

605
17.2.77



Psychological Review

CHARLES N. COFER, EDITOR
Pennsylvania State University

VOLUME 72, 1965

HELEN ORR
Managing Editor

ELIZABETH S. REED
Advertising Manager

VIRGINIA WRIGHT
Editorial Assistant

PUBLISHED BIMONTHLY BY THE
AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
· PRINCE AND LEMON STS., LANCASTER, PA. 17604
AND 1200 SEVENTEENTH ST. N. W., WASHINGTON, D. C. 20036

Editorial Consultants

JACK A. ADAMS
 R. B. AMMONS
 NORMAN H. ANDERSON
 MORTIMER H. APLEY
 SOLOMON E. ASCH
 RICHARD C. ATKINSON
 RICHARD S. BALVIN
 HAROLD P. BECHTOLDT
 IRWIN A. BERG
 EDWARD A. BILODEAU
 ROBERT R. BLAKE
 GORDON H. BOWER
 ROBERT M. BOYNTON
 F. ROBERT BRUSH
 C. J. BURKE
 ROGER V. BURTON
 ROBERT R. BUSH
 ARNOLD H. BUSS
 DONALD T. CAMPBELL
 EDWARD C. CARTERETTE
 RAYMOND B. CATTELL
 RUSSELL M. CHURCH
 JACOB COHEN
 TOM N. CORNSWEET
 WALTER HOBSON CROCKETT
 LEE J. CRONBACH
 M. R. D'AMATO
 SOLOMON DIAMOND
 JAMES C. DIGGORY
 ALLEN T. DITTMANN
 ELIZABETH DUFFY
 JAMES P. EGAN
 DAVID EHRENFREUND
 HENRY C. ELLIS
 TRYGG ENGEN
 SUSAN ERVIN-TRIPP
 PAUL M. FITTS
 ROBERT B. FREEMAN, JR.
 BENJAMIN FRUCHTER
 ROBERT M. GAGNE
 WENDELL R. GARNER
 HAROLD B. GERARD

MURRAY GLANZER
 ISRAEL GOLDIAMDOND
 ALBERT E. GOSS
 CLARENCE H. GRAHAM
 JAMES J. GIBSON
 G. ROBERT GRICE
 J. P. GUILFORD
 FRITZ HEIDER
 HARRY HELSON
 JERRY HIRSCH
 JULIAN HOCHBERG
 HOWARD S. HOFFMANN
 ROBERT R. HOLT
 LLOYD G. HUMPHREYS
 EARL B. HUNT
 HOWARD F. HUNT
 LEO M. HURVICH
 IRVING L. JANIS
 MURRAY E. JARVIK
 JAMES J. JENKINS
 ARTHUR R. JENSEN
 NEHEMIAH JORDAN
 HENRY F. KAISER
 TRACY S. KENDLER
 JOHN L. KENNEDY
 HERBERT D. KIMMEL
 DOUGLAS H. LAWRENCE
 DANIEL S. LEHRMAN
 ALVIN M. LIBERMAN
 FREDERIC M. LORD
 MAURICE LORR
 R. DUNCAN LUCE
 KENNETH MACCORQUODALE
 ROBERT B. MACLEOD
 GEORGE F. MAHL
 ROBERT B. MALMO
 GEORGE MANDLER
 ROBERT A. MCCLEARY
 WILLIAM J. MCGUIRE
 ARNOLD MECHANIC
 ARTHUR W. MELTON

NEAL E. MILLER
 PETER M. MILNER
 HOWARD MOLTZ
 CLIFFORD T. MORGAN
 BENNET B. MURDOCK, JR.
 ULRIC NEISSER
 DAVID S. PALERMO
 IRWIN POLLACK
 WILLIAM F. PROKASY
 FRANK RESTLE
 DONALD A. RILEY
 IRVIN ROCK
 DAVID L. ROSENHAN
 PHILIP J. RUNKEL
 WALLACE A. RUSSELL
 THEODORE R. SARBIN
 LOWELL M. SCHIPPER
 HAROLD SCHLOSBERG
 DAVID SHAKOW
 JEROME E. SINGER
 NORMAN J. SLAMECKA
 OLIN W. SMITH
 RICHARD L. SOLOMON
 JOHN S. STAMM
 JULIAN C. STANLEY
 ELIOT STELLAR
 S. SMITH STEVENS
 SAMUEL SUTTON
 RENATO TAGIURI
 PHILIP TEITELBAUM
 HANS-LUKAS TEUBER
 ENDEL TULVING
 BENTON J. UNDERWOOD
 SEYMOUR WAPNER
 J. M. WARREN
 JOSEPH WEITZ
 MICHAEL WERTHEIMER
 JOACHIM F. WOHLWILL
 DAVID ZEAMAN
 ROBERT R. ZIMMERMANN

CONTENTS OF VOLUME 72

BAHRICK, HARRY P. The Ebb of Retention.....	60
BEVER, T. G., FODOR, J. A., AND WEKSEL, W. Is Linguistics Empirical?.....	493
BEVER, T. G., FODOR, J. A., AND WEKSEL, W. On the Acquisition of Syntax: A Critique of "Contextual Generalization".....	467
BLACK, ROGER W. On the Combination of Drive and Incentive Motivation.....	310
BOGARTZ, RICHARD S. On the Assumption of a Steeper Avoidance Gradient in Miller's Conflict Theory.....	162
BOURNE, LYLE E., JR. See HAYGOOD, ROBERT C.	
BRAINE, MARTIN D. S. On the Basis of Phrase Structure: A Reply to Bever, Fodor, and Weksel.....	483
BUGELSKI, B. R. In Defense of Remote Associations.....	169
CANON, LANCE KIRKPATRICK. See FESTINGER, LEON.	
COFER, CHARLES N. Editorial.....	1
CORBALLIS, M. C. Practice and the Simplex.....	399
DALLETT, KENT M. In Defense of Remote Associations.....	164
DAY, R. H., AND POWER, R. P. Apparent Reversal (Oscillation) of Rotary Motion in Depth: An Investigation and a General Theory.....	117
FESTINGER, LEON, AND CANON, LANCE KIRKPATRICK. Information about Spatial Location Based on Knowledge about Efference.....	373
FLOCK, HOWARD R. Optical Texture and Linear Perspective as Stimuli for Slant Perception	505
FOA, URIEL G. New Developments in Facet Design and Analysis.....	262
FODOR, J. A. See BEVER, T. G.	
FREEMAN, ROBERT B., JR. Ecological Optics and Visual Slant.....	501
HAMMOND, KENNETH R., AND SUMMERS, DAVID A. Cognitive Dependence on Linear and Nonlinear Cues.....	215
HARRIS, CHARLES SAMUEL. Perceptual Adaptation to Inverted, Reversed, and Displaced Vision.....	419
HAYGOOD, ROBERT C., AND BOURNE, LYLE E., JR. Attribute- and Rule-Learning Aspects of Conceptual Behavior.....	175
JONES, FRANK PIERCE. Method for Changing Stereotyped Response Patterns by the Inhibition of Certain Postural Sets.....	196
JUNG, JOHN. Comments on Mandler's "From Association to Structure".....	318
LANE, HARLAN. The Motor Theory of Speech Perception: A Critical Review.....	275
LEWIS, DONALD J., AND MAHER, BRENDAN A. Neural Consolidation and Electroconvulsive Shock.....	225
LOEHLIN, JOHN C. Some Methodological Problems in Cattell's Multiple Abstract Variance Analysis.....	156
LOEVINGER, JANE. Person and Population as Psychometric Concepts.....	143
MAHER, BRENDAN A. See LEWIS, DONALD J.	
MALMO, ROBERT B. Comment on the Exchange of Theoretical Notes between Smith and Black and Lang.....	240
MANDLER, GEORGE. Subjects Do Think: A Reply to Jung's Comments.....	323
MARTIN, EDWIN. Transfer of Verbal Paired Associates.....	327
MINARD, JAMES G. Response-Bias Interpretation of "Perceptual Defense": A Selective Review and Evaluation of Recent Research.....	74

MOLTZ, HOWARD. Contemporary Instinct Theory and the Fixed Action Pattern....	27
NOLAND, J. H. See SCOTT, THOMAS R.	
NORMAN, DONALD A. See WAUGH, NANCY C.	
O'DONOVAN, DENIS. Rating Extremity: Pathology or Meaningfulness?.....	358
PARDUCCI, ALLEN. Category Judgment: A Range-Frequency Model.....	407
PETERSON, DONALD R. Scope and Generality of Verbally Defined Personality Factors	48
POWER, R. P. See DAY, R. H.	
REYNOLDS, ROBERT W. An Irritative Hypothesis Concerning the Hypothalamic Regulation of Food Intake.....	105
RINN, JOHN L. Structure of Phenomenal Domains.....	445
ROKEACH, MILTON, AND ROTHMAN, GILBERT. The Principle of Belief Congruence and the Congruity Principle as Models of Cognitive Interaction.....	128
ROTHMAN, GILBERT. See Rokeach, Milton.	
SCHARF, BERTRAM. See ZWICKER, EBERHARD.	
SCOTT, THOMAS R., AND NOLAND, J. H. Some Stimulus Dimensions of Rotating Spirals	344
SLAMECKA, NORMAN J. In Defense of a New Approach to Old Phenomena.....	242
STOLLNITZ, FRED. Spatial Variables, Observing Responses, and Discrimination Learning Sets.....	247
SUMMERS, DAVID A. See HAMMOND, KENNETH R.	
THOMPSON, ROBERT. Centrencephalic Theory and Interhemispheric Transfer of Visual Habits.....	385
WAUGH, NANCY C., AND NORMAN, DONALD A. Primary Memory.....	89
WEKSEL, W. See BEVER, T. G.	
ZWICKER, EBERHARD, AND SCHARF, BERTRAM. A Model of Loudness Summation....	3

PSYCHOLOGICAL REVIEW

EDITORIAL

Richard Solomon's distinguished editorship of the *Review* is being completed in the present issue.¹ However, the processing of manuscripts was transferred to the new Editor in January 1964. On the basis of a year's experience with manuscripts submitted for publication in the *Review*, certain comments seem to be in order concerning the policy as to what kinds of articles are appropriate for the *Review*.

The *Review* is the APA journal concerned with psychological theory and its development. The guiding principle used by the Editor and his consultants in evaluating manuscripts for their *appropriateness* to the *Review* is that the major theme of an article must be a theoretical one (in any area of psychology). Many manuscripts received in the past year have been returned to their authors because they do not meet the criterion of appropriateness. Such manuscripts fall into several classes.

1. *Literature Reviews*. Surveys and summaries of the literature on some topic or subfield are not appropriate for the *Review*. They are the province of the *Psychological Bulletin*. A theoretical article may, of course, often be supported by a review of the literature. The critical feature of such an article, however, if it is to be appropriate for the *Review*, must be a theory

or a theoretical development. Reviews of the literature may lead to rejection of or the discovery of critical weaknesses in a theoretical position, but in the preparation of such papers the theoretical question should be paramount and the statement of a theoretical development or alternative should receive careful and detailed attention.

2. *Statistical Articles*. Papers on statistics and problems of experimental design are again the province of the *Bulletin*. (It should be noted that the *Bulletin* encourages interpretive articles. A *Bulletin* article should serve as a bridge between statistical contributions per se and the typical research psychologist.)

3. *Methodological Articles*. These, too, belong to the *Bulletin*, unless a flaw in method has led to theoretical misconceptions of major degree. Papers concerned with the use of control groups, specific pieces of apparatus, comparisons of methods, and the like are not appropriate for the *Review*.

4. *Empirical Studies*. Reports of experiments and other kinds of studies as such are not appropriate for the *Review*, even though they test a hypothesis or a deduction from some theory. Empirical studies which in-validate a theory or hypothesis are usually not appropriate, either, unless the major burden of the paper is a theoretical development. It is a time-honored practice for empirical journals to publish studies designed to test hy-

¹ The papers by Minard, Bahrack, Moltz, and Peterson, published in this issue, were processed by Richard Solomon.

potheses, and this is true no matter whether the outcome is favorable or unfavorable to the hypothesis or deduction.

Empirical work may be appropriately included in an article for the *Review*, but only when the report of empirical material occurs in the context of theoretical matters. The major burden of the paper must be theoretical, with the empirical investigation in a supporting role. The report of the study or experiment must be sufficiently detailed that it can be evaluated by the criteria used in empirical journals.

5. *Papers Expressing an Idea, a Definition or a Concept.* A number of papers have been received whose content is "thin." They may include an idea, or concept, or perhaps a definition, but do not reflect a detailed working out of the theoretical aspect or implications. Such papers do not meet the standards of a *Psychological Review* article. Such papers, however, if brief, may be publishable in the Theoretical Notes section, which will continue under the new Editor, and should be prepared with this goal.

6. *Papers Prepared for Other Purposes.* Speeches, symposium presentations, technical reports, and term

papers have been submitted for publication in unaltered and unedited form. While such papers may be good ones, their form suggests that the authors have not considered very seriously the nature of the journal and of its audience to which they have sent the papers. Papers of this kind are seldom publishable in the *Review* as they are submitted and may be returned to the authors summarily without editorial review. All manuscripts submitted must conform to the usages of the *APA Publication Manual*.

There are, of course, papers whose appropriateness for the *Review* is difficult to decide on the basis of the guidelines just indicated. Authors, in such cases, are free to send their manuscripts to the journal they deem to be appropriate. It usually serves no purpose to write to the Editor for an opinion of the appropriateness of a paper described in a general way.

The new Editor has continued and will continue the practice of using a large number of consultants, each one only on occasion, to assist him in the task of evaluating manuscripts. To those who have already served and to those who will serve in the future, he extends grateful acknowledgment of their assistance.

A MODEL OF LOUDNESS SUMMATION¹

EBERHARD ZWICKER

Technische Hochschule Stuttgart

BERTRAM SCHARF

AND

Northeastern University

A psychophysical model is presented that explains why loudness summates across frequency as it does and that permits the precise calculation of loudness from the physical spectrum. Loudness is represented by geometrical patterns derived from the masking of pure tones by narrow bands of noise. The masking patterns are converted to loudness patterns by means of the critical-band function that relates tonalness in Barks to frequency in cycles per second and a power function that relates specific loudness, loudness per Bark, to sound pressure level (SPL). Plotted on the coordinates of specific loudness and tonalness, the geometrical patterns are integrated to yield a value in sones₀ for the overall loudness. Calculated values are compared to experimental values obtained from loudness balances with 3 types of sound.

Loudness summates across frequency. Thus two pure tones are usually louder than either alone. Their overall loudness is seldom, however, the sum of their individual loudnesses. The failure of loudness to summate perfectly is ascribed to mutual inhibition between the tones (cf. Stevens, 1956). Since the degree of inhibition depends upon both the intensity of the tones and their frequency separation, no simple rule can serve to predict their total loudness directly from their physical properties. Nevertheless, it is possible to represent each of the two tones, or indeed any number of tones including the infinite number contained in a continuous spectrum, by geometrical patterns that permit a precise analysis of loudness summation. These patterns are the essential feature of the model of loudness summation presented in this paper.

Although often described below in broad physiological terms, the model is based entirely on psychophysical measures of masking, loudness, and the critical band, and is therefore wholly psychological. The scales of "specific loudness" and "tonalness," which replace the physical scales of intensity and frequency, provide the coordinates for the geometrical representation of loudness. The specific loudness scale, unlike the sone scale, is a derived scale and represents the loudness produced over each unit of tonalness. The scale of tonalness is based on the critical band as measured in experiments on the threshold for complex sounds, two-tone masking, phase sensitivity, and loudness summation itself (Feldtkeller, 1955; Scharf, 1961b).

The use of a geometrical representation of loudness follows from the place theory of hearing which assumes a local pattern of excitation produced within the auditory nervous system by a pure tone or by each pure tone component of a complex sound (a sound composed of two or more frequency components). In the present model, the shape and level of the local patterns

¹ Preparation of this paper was supported in part by United States Public Health Service Research Grant NB 04464 from the National Institute of Neurological Diseases and Blindness. The authors wish to thank S. S. Stevens for helpful discussions and suggestions concerning the paper.

are inferred from the masking of pure tones by narrow bands of noise. A complex sound is represented by a number of these local patterns whose integrated area yields a measure of the total loudness.

Some of the basic notions underlying the model such as the conversion of the intensity and frequency scales to psychological scales and the derivation of loudness patterns from masking patterns are not new (Fletcher & Munson, 1937; Harris, 1959; Howes, 1950; Munson & Gardner, 1950). New are the concept of the specific loudness and the use of the critical band and also the application of the model to a number of different types of sound.

Parts of the model have already been described elsewhere (Scharf, 1964; Zwicker, 1958, 1963). The whole model is here considered and tested against a variety of empirical measurements of loudness. As will be seen, the model, based on a few simple and general rules, accounts for most of the complex data on loudness summation.

Before introducing the model and testing it against experimental measures, we review briefly the basic facts of loudness summation.

BASIC FACTS OF LOUDNESS SUMMATION

Perhaps the most striking fact of loudness summation, defined as the increase in loudness with the spread of energy over frequency, is that it does not begin until the energy is spread over more than a critical bandwidth (*Frequenzgruppe*) (Zwicker & Feldtkeller, 1955; Zwicker, Flottorp, & Stevens, 1957). Up to the critical band, the loudness of a complex sound is largely independent of the width of energy spread. Loudness first and rather abruptly begins to increase with

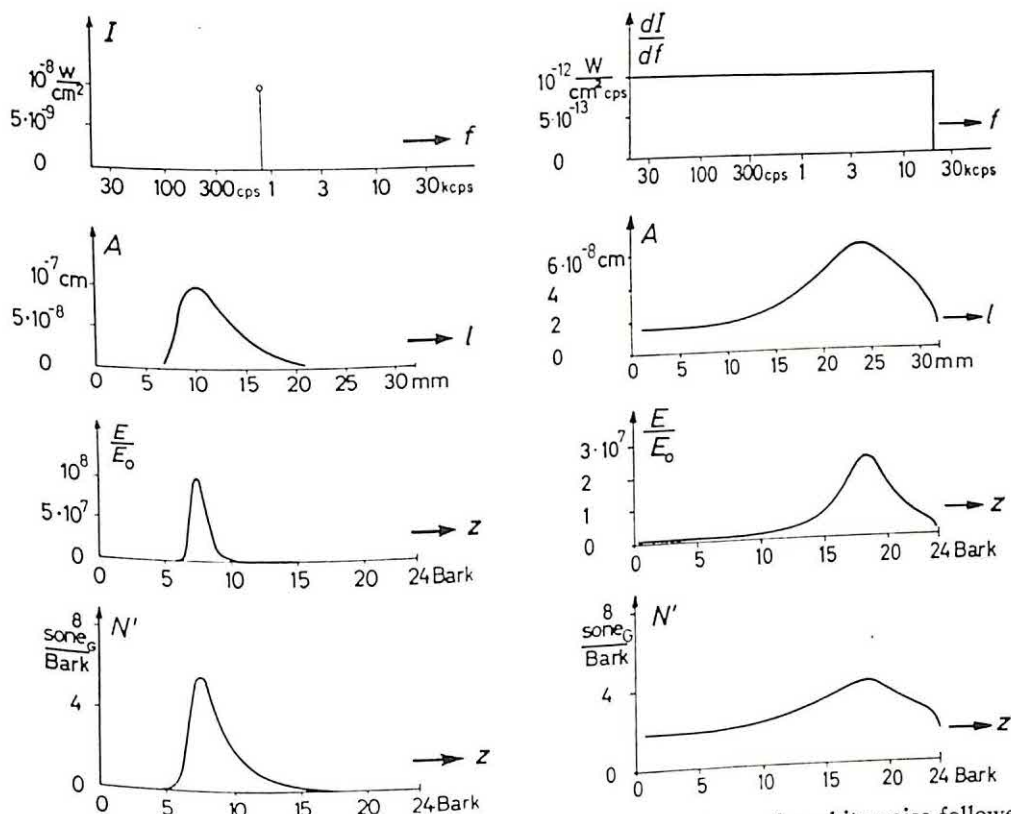
bandwidth at the critical band. The value of a critical band centered on a given frequency is the same at all intensities, but beyond it the amount of the increase in loudness depends upon the intensity of the sound (Zwicker & Feldtkeller, 1955; Zwicker, Flottorp, & Stevens, 1957). Indeed, close to threshold the loudness decreases slightly rather than increases (Scharf, 1959a). At sensation levels above 10 decibels, however, some increase in loudness with bandwidth may be noted, with the greatest increases at moderate sensation levels between 40 and 60 decibels.

These rules hold regardless of where in the frequency spectrum the sound is located (Zwicker & Feldtkeller, 1955; Scharf, 1959a). They also hold whether the threshold is normal or is raised to an abnormally high level by the presence of a moderately intense uniform masking noise (Scharf, 1961a) or by a middle-ear impairment (Scharf, 1962b). The complex sound itself may be composed of any number of components from two up to the infinite number comprising a band of white noise (Scharf, 1959b). A complex sound at a given sensation level is loudest, however, when the components are all equally loud (Scharf, 1962a) and when they are separated by an equal number of critical bands (Zwicker, Flottorp, & Stevens, 1957).

The relation between loudness and bandwidth is the same even for sounds of short duration, although the loudness of a complex sound decreases as its duration is decreased below 70 milliseconds (Port, 1963a, 1963b).

DESCRIPTION AND DEVELOPMENT OF THE MODEL

Although the experimental measures of loudness summation yield a complex pattern of results, a single model



1a. First diagram is the physical spectrum of an 800-cps tone. (Intensity I plotted against frequency f . Next is plotted the Amplitude A of the displacement caused by the tone on the basilar membrane as a function of Distance l from the helicotrema. Third is shown the Excitation E at the Organ of Corti as a function of tonalness z . [The symbol E_0 is an arbitrary reference.] Last is shown the Specific Loudness N' as a function of Tonalness z . The area under the loudness pattern corresponds to the total loudness of the tone.)

1b. The spectrum of a white noise followed by the displacement, excitation, and loudness patterns produced by the noise (Zwicker, 1958, adapted with permission of *Acustica*).

FIG. 1. Sequential formation of a loudness pattern.

can encompass them all. First let us outline briefly the model as it is applied to a single tone and to white noise. Then we shall review the detailed development of the model.

Figure 1 traces the transformations that are thought to occur when the ear is stimulated by a pure tone (Fig. 1a) and by white noise (Fig. 1b). The physical measurement of an ideal pure-

tone stimulus at the ear drum yields a line spectrum that shows all the sound energy concentrated at a single frequency. Beyond the ear drum, the tonal intensity and frequency are transmitted by the ossicles of the middle ear to the oval window. At that point the inner-ear fluid and consequently the basilar membrane are set in motion. Békésy's measures of the vibra-

tion of the basilar membrane permit us to plot the maximum amplitude A of the displacement produced by the tone over the length of the basilar membrane, as shown in the second part of Figure 1a. This first, important step in the development of a model for loudness summation shows that energy confined to a single point on the physical frequency scale produces a displacement over a wide area of the basilar membrane. The displacement of the basilar membrane somehow activates the sensory cells whose neural response is then translated through several synaptic junctures to higher levels of the nervous system. The wide lateral displacement on the basilar membrane is apparently greatly reduced at the neural level, and the pattern of Excitation E looks like that in the third part of the figure. (The excitation pattern is based primarily upon masking data, as explained below.)

The value of the excitation relative to an arbitrary reference E_0 is plotted here against the scale of tonalness (z) which is related in an approximately linear fashion to distance on the basilar membrane. The excitation pattern is then transformed into a loudness pattern, depicted in the fourth part of Figure 1a, by means of a special equation relating Specific Loudness N' to Excitation E . The total loudness produced by the original tone is the integral of the area under the specific loudness pattern.

Thus the first part of the figure comes from a purely physical measure. The second part, the displacement over the basilar membrane, comes from physiological measures. The last two parts of the figure may, in the next years or decades, also be measured physiologically, but in the meantime they are derived from psychophysical measures to form the basis of our

model of loudness and loudness summation.

Now let us look at the same sequence for white noise. The energy in white noise is distributed over frequency as shown in the first part of Figure 1b. The displacement produced along the basilar membrane is approximated from measurements on models of the basilar membrane (Oettinger & Häuser, 1961). The excitation and loudness patterns for the noise are quite different from the patterns produced by a pure tone.

Whereas the first two diagrams of Figure 1 are well-known physical and physiological measures, the representations of excitation and specific loudness are new and require explanation. The explanation may be broken down into five parts to which a sixth is added for sounds heard against a background of noise.

1. The source of the excitation measures
2. The relation between the tonalness scale and the physical frequency scale
3. The function relating the specific loudness to excitation
4. The dependence of the transmission characteristics of the middle ear on frequency
5. The generation of the total loudness
6. The effect of partial masking on loudness

1. It is assumed that both the excitation and the loudness patterns generated in the nervous system upon acoustic stimulation can be derived from the masking patterns, which can be empirically measured. This assumption is similar to that originally set forth by Fletcher and Munson (1937) and adopted by Steinberg and Gardner (1937) and also by Harris

(1959). The masking pattern for a given sound is obtained by using the sound to mask a pure tone. The plot of the masked threshold for the tone, as a function of its frequency, is the masking pattern.

In order to avoid the necessity of measuring the masked thresholds with each sound to which the model is applied, a set of standard masking patterns and their derived excitation patterns have been made available for narrow bands of noise (Zwicker, 1958, 1963). These narrow-band patterns represent the masking patterns of any sound narrower than or equal to a critical band, including pure tones, since the masking patterns are essentially the same for all bandwidths less than a critical band (Ehmer, 1959; Zwicker, 1956). Combinations of patterns centered at different frequencies can represent the masking patterns of sounds wider than a critical band. A set of masking patterns is shown in Figure 2 for the masking of pure tones by a narrow band of noise centered at 1,200 cps. The parameter on the curves is the sound pressure level (SPL) of the masking noise.

The first step in the transformation from masking patterns to excitation patterns is the conversion of the masking level at each frequency to the corresponding excitation level. But how do we determine from the measured masked threshold the excitation level within the nervous system produced by the masking sound? The answer comes from data on the self-masking of noise, i.e., from measures of the jnd for intensity. Since in such measures the excitation patterns of the masked noise, ΔI , and the masking noise, I , are almost identical, their ratio, $\Delta I/I$, may be taken to express the ratio between the excitation of the just masked stimulus and that of the masking stimulus.

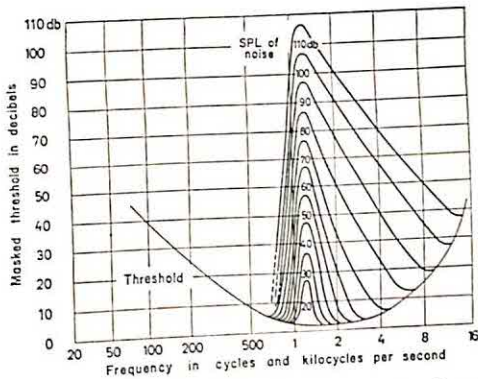


FIG. 2. Masking effect of a narrow band of noise with a center frequency of 1,200 cps. The parameter is the effective SPL of the noise. (On the ordinate the masked threshold for a pure tone is plotted for each level of noise as a function of frequency [Zwicker, 1958, adapted with permission of *Acustica*].)

Fortunately over a wide range of intensities the ratio, $\Delta I/I$, is fairly constant, being equal to $\frac{1}{2}$ for the lower and middle frequencies and to $\frac{1}{4}$ for the higher frequencies (Zwicker, 1956). The minimum excitation required then to mask completely at a given frequency is twice (3 decibels) or four times (6 decibels) the intensity at the masked threshold. The addition of 3 (or 6) decibels to the masked threshold gives the value for the excitation level. Excitation level will be expressed in decibels as $L_E = 10 \log E/E_0$ where E_0 is a reference value corresponding to $I_0 = 10^{-16}$ watt/centimeters².

2. The representation of the excitation pattern requires not only the conversion of the ordinate from masking level to excitation level, but also the conversion of the abscissa from the physical frequency scale to one that corresponds more closely to the way in which activity spreads in the auditory system. A good correspondence appears to be achieved by the tonalness or z scale. (The spread of activity represented by the z scale probably corresponds to a spatial spread of activity

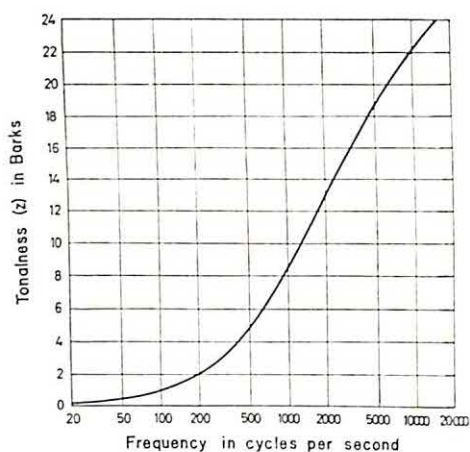


FIG. 3. Relation between tonalness in Barks and frequency (Zwicker, 1961, adapted with permission of *Journal of the Acoustical Society of America*).

on the basilar membrane, but at higher neural levels, the spread is not necessarily spatial.) The relation between tonalness and frequency is given by the function in Figure 3 (Zwicker, 1961). The unit, Bark (after the German acoustician Barkhausen), replaces the critical band in order to distinguish the Bark as a measure of sensory events from the critical band, a stimulus measure. The z function seems to be an appropriate measure of auditory activity because of its crucial role as the unit of integration in loudness summation, threshold measurements, and masking (Greenwood, 1961; Scharf, 1961b). Moreover, a similar function relates the jnd of pitch, the frequency representation on the basilar membrane, and the mel scale of pitch to frequency.

With the aid of the tonalness function, the masking patterns of Figure 2 have been converted to excitation patterns in Figure 4. (Excitation patterns for narrow bands of noise centered at lower and higher frequencies are similar to those in Figure 4; see Zwicker, 1958.) The excitation patterns are for

a band of noise no wider than a critical band, centered at 1,200 cps and having the SPL shown as the parameter on the curve. These curves also represent the excitation produced by a 1,200-cps tone, since the masking patterns produced by pure tones and by narrow bands of noise are almost the same (except where beats and audible non-linear distortions occur).

3. In order to use these excitation patterns for the derivation of loudness, it is necessary to know the relation between loudness and excitation or more exactly, between Specific Loudness N' and Excitation E . A psychophysical equation expressing this relation has been formulated (Zwicker, 1958, 1963). It is based upon Stevens' power law, which says in one form that equal intensity ratios yield equal loudness ratios (Stevens, 1957); it is also based upon the assumption that the loudness of any sound is the integral of the specific loudness over the z scale. In this sense every sound involves the summation of loudness, for Figure 4 shows that the excitation aroused by even a pure tone spreads over a considerable portion of the z scale. The next few paragraphs trace the deriva-

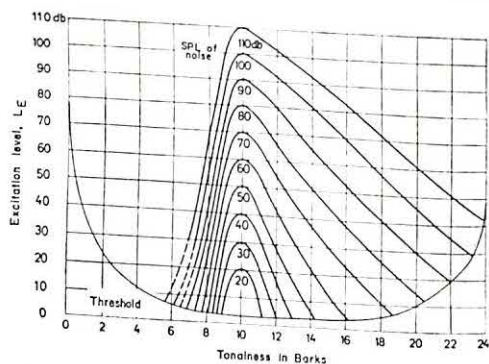


FIG. 4. Excitation patterns calculated from the masking produced by a narrow band of noise centered at 1,200 cps. (The level of the effective SPL of the band of noise is the parameter on the curves, Zwicker, 1958; adapted with permission of *Acustica*.)

tion of the basic formula that relates specific loudness to excitation.

Stevens' law may be expressed as:

$$\frac{\Delta N}{N} = k' \frac{\Delta I}{I} \quad [1]$$

where I is the intensity of a tone and N is the loudness. In terms of excitation, we must assume that this law applies only to the excitation over a single Bark. It is not feasible to apply this law to the whole excitation produced by a pure tone since the excitation pattern changes considerably both with frequency and intensity. The critical band or Bark provides a constant unit over which the excitation can be converted to loudness by means of a single equation. The loudness due to the excitation over a single Bark is called the "specific loudness," represented by the symbol N' . With Intensity I replaced by Excitation E , Equation 1 becomes:

$$\frac{\Delta N'}{N'} = k \frac{\Delta E}{E}. \quad [2]$$

Near threshold, where intensity discrimination is poor, the relationship breaks down unless a constant is added to the denominators of Equation 2. This constant, E_{gr} , may be thought to represent the excitation produced in the ear by an inaudible physiological background noise. This excitation can suppress a weak excitation produced by an external stimulus thereby setting a lower limit, the absolute threshold, for the ear's sensitivity. The corresponding inaudible specific "loudness" is N'_{gr} .

$$\frac{\Delta N'}{N' + N'_{gr}} = k \frac{\Delta E}{E + E_{gr}} \quad [3]$$

Treating Equation 3 as a differential equation and integrating, we have

$$\log (N'_{gr} + N') = k \log (E_{gr} + E) + \log C \text{ or } [4]$$

$$N'_{gr} + N' = C (E_{gr} + E)^k. \quad [4a]$$

Taking as our boundary conditions $N' = 0$ when $E = 0$, we obtain for the constant of Integration C :

$$C = \frac{N'_{gr}}{E_{gr}^k} \text{ or } [5]$$

$$N'_{gr} = C (E_{gr})^k. \quad [5a]$$

The evaluation of the constant E_{gr} depends upon the same assumption used to convert from masking to excitation patterns, namely, that the masking excitation must be twice or four times the excitation produced by the just masked tone. It is assumed that the internal background excitation is twice the excitation E_t produced by an external tone at the absolute threshold (or four times if the tone is at a high frequency).

$$E_{gr} = 2E_t \quad [6]$$

Using Equations 5 and 6 to substitute in Equation 4 we arrive at:

$$N' = N'_{gr} \left[\left(\frac{E}{2E_t} + 1 \right)^k - 1 \right]. \quad [7]$$

In order to express the value N'_{gr} in relative values of Excitation E_t/E_o , we introduce the reference value N'_{gr_o} . Using Equation 5a with $2E_t$ replacing E_{gr} , we obtain:

$$\frac{N'_{gr}}{N'_{gr_o}} = \left(\frac{2E_t}{E_o} \right)^k \text{ or } [8]$$

$$N'_{gr} = N'_{gr_o} \left(\frac{2E_t}{E_o} \right)^k.$$

Equation 7 may now be written:

$$N' = N'_{gr_o} \left(\frac{2E_t}{E_o} \right)^k \times \left[\left(\frac{E}{2E_t} + 1 \right)^k - 1 \right]. \quad [9]$$

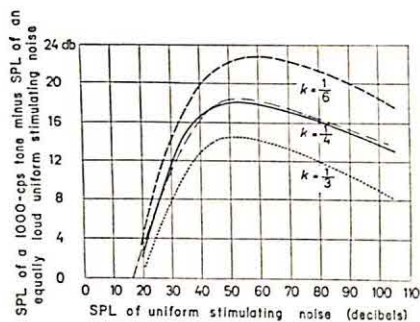


FIG. 5. Calculated and observed (dot-dash line) difference in SPL between a 1,000-cps tone and a uniform masking noise. (Calculations were made with three different values of k inserted in Equation 9 [Zwicker, 1958; adapted with permission of *Acustica*].)

N'_{gr_0} is an arbitrary constant that determines the units in which the specific loudness is measured; E_0 is the excitation corresponding to I_0 taken as 10^{-16} watt/centimeters²; E is the excitation, at a given point on the z scale, evoked by the stimulating tone.

The exponent k remains to be determined. Its importance becomes clear at large values of E where Equation 9 approaches the simpler relation $N' \approx (E/E_t)^k$. The value of k can not, however, be determined from Equation 9 by the direct measurement of N' , the loudness arising from excitation restricted to a single Bark. Such a restricted excitation pattern can not be produced in the ear. This difficulty has been bypassed, and the value of the exponent calculated from measurements of the loudness of two very different sounds (Zwicker, 1958). One sound was a 1,000-cps tone and the other a uniform stimulating noise, a noise that has the same overall intensity in each critical band. The two sounds were matched for loudness at various levels. The observed differences in loudness level between the noise and tone are plotted as the dot-dash line in Figure 5. On the same plot are shown the differences calcu-

lated from Equation 9 with 3 different values of k . In these calculations the Total Loudness N of the tone and also of the noise is the integral of the specific loudness over the z scale, that is

$$N = \int_{z=0}^{z=24} N' dz. \quad [10]$$

The best fit between the observed and calculated differences is obtained with $k = 0.23$.

All the unknown constants of Equation 9 have now been determined except the arbitrary constant N'_{gr_0} . This constant is chosen so that the integral, $\int N' dz$, of the specific loudness of a 1,000-cps tone at 40-decibel SPL is equal to 1 sone. The unit assigned to the specific loudness is sone_G, where the subscript G distinguishes this calculated value from the sone obtained by direct loudness measurements.

After some smoothing at low levels, we arrive at the final equation for the specific loudness.

$$N' = 0.08 \left(\frac{E_t}{E_0} \right)^{.23} \times \left[\left(1/2 \frac{E}{E_t} + 1/2 \right)^{.23} - 1 \right] \text{ in sone}_G/\text{Bark} \quad [11]$$

This equation is graphed in Figure 6 with both E_t and E expressed in decibels as L_t and L_E . Given the excitation level at threshold and given the excitation level due to the stimulating tone, the specific loudness, N' in sones_G, may be found on the graph at each Bark under the excitation pattern.

4. The ordinate of Figure 4 should be labeled not L_E , but $L_E - a$. The symbol a corrects the error introduced by the fact that the masked thresholds for pure tones vary not only because the excitation levels produced within the nervous system by a masking stimulus change with frequency, but also

because more energy is transmitted by the middle ear to the receptor cells at some frequencies than at others. The value of a is in a sense the amount of stimulus energy lost (or gained, depending on the reference value chosen) as a given tone is transmitted to the oval window. The masked threshold for the tone would then be greater by a decibels and would yield a value for the excitation level L_E of the masking stimulus that would be a decibels too high.

Direct measures of the transmission characteristics of the middle ear are apparently not available, but they may be approximated if we assume that the transmission characteristics govern the dependence of the absolute threshold on frequency (at least above 2,000 cps). Such an assumption implies that neural sensitivity is constant over most of the audible frequency scale (up to 9,000 or 10,000 cps) and that, for example, the average threshold is 18 decibels lower at 4,000 cps than at 8,000 cps (Robinson, 1957) because the energy transmitted to the oval window is 18 decibels less at 8,000 cps. The value of a would then be 18 decibels higher at 8,000 cps than at 4,000 cps.

Below 2,000 cps, the rise in the absolute threshold with decreasing frequency is ascribed to increased internal noise rather than to reduced middle-ear transmission. Since the internal noise has already been included in Equation 9 as part of the expression E_t/E_o , no other account need be taken of it here. Reference to a is omitted in the rest of this paper for the sake of simplicity.

5. The model of loudness described so far for a single tone or narrow band of noise may also be applied to complex sounds such as wide bands of noise and multiple tones. Two alternative procedures are available.

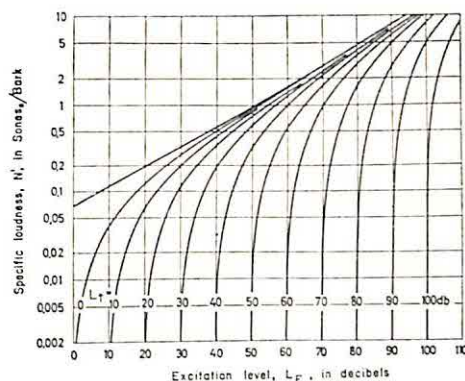


FIG. 6. Specific Loudness N' in sones/Bark as a function of the Excitation Level L_E . The exponent is $k = \frac{1}{4}$. (The parameter on the curves is the Excitation Level L_t of the test tone at threshold [Zwicker, 1958; adapted with permission of *Acustica*].)

The first procedure is exactly like that described above for obtaining the loudness value for a single tone. The masking pattern produced by the complex, as masking sound, is measured and transformed to the excitation pattern. The specific loudness can be read from Figure 6 as a function of tonalness to yield a loudness pattern whose integral over z is the total loudness. In the second procedure, the measurement of the masking pattern of the complex sound is replaced by the much simpler measurement of the SPL in each component critical band of the stimulus. Each band is treated as an independent stimulus that gives rise to an excitation pattern like those shown in Figure 4. The overlapping patterns thus obtained have a common upper envelope, which determines the excitation level; wherever two or more patterns overlap, only the highest excitation level is used. This procedure requires, of course, the prior measurement of the masking pattern for bands of noise at various SPLs and centered at various frequencies. Fortunately the shape of the masking patterns changes slowly enough with center fre-

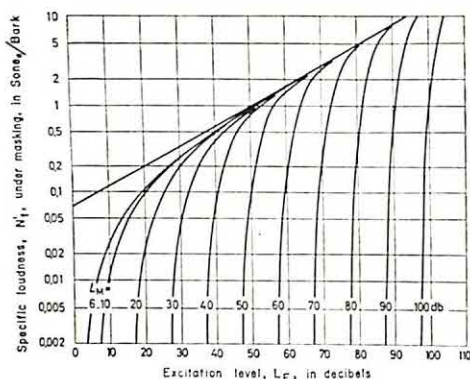


FIG. 7. Specific Loudness N'_t of partially masked pure tones as a function of the excitation level of the tone. (The parameter on the curves is the Excitation Level L_M of the masking sound [Zwicker, 1963; adapted with permission of *Acustica*].)

quency so that measurements of the masking patterns at only 6 or 7 scattered frequencies suffice to give good approximations of the patterns over the whole audible frequency spectrum.

Examples of the application of both procedures are given below.

6. The loudness of partially masked sounds must be treated somewhat differently from the loudness of unmasked sounds. A sound is partially masked when its loudness is reduced by the presence of another sound. Partial masking occurs (and the measurement of the usual masked thresholds is possible) because the components of a complex sound may be distinguished from one another; their loudnesses do not always summate to yield an overall loudness. The question now posed is how our model can handle partial masking.

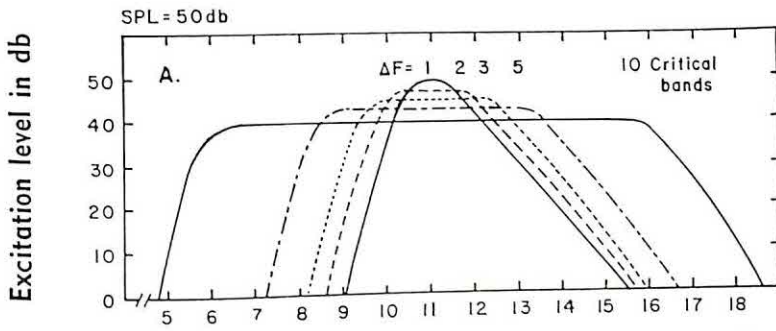
A series of measurements (Zwicker, 1963) showed that the present model underestimates the loudness of a 1,000-cps tone partially masked by either narrow-band or wide-band noise at various levels. This underestimation could be corrected, however, by an adjustment of the curves in Figure 6.

Apparently when the excitation at threshold, L_t , is produced by an external masking noise, the loudness of a partially masked tone increases more rapidly with intensity than when the threshold excitation is produced by internal noise. Consequently, the curves of Figure 6 must be steeper if they are to be used with partially masked sounds. Adjustment of the curves in Figure 6 to yield loudness patterns for a partially masked tone whose integral better approximates the measured loudness leads to the curves of Figure 7. In place of the excitation level, L_t , at threshold, the excitation level, L_M , of the masking stimulus is the parameter on the curves.

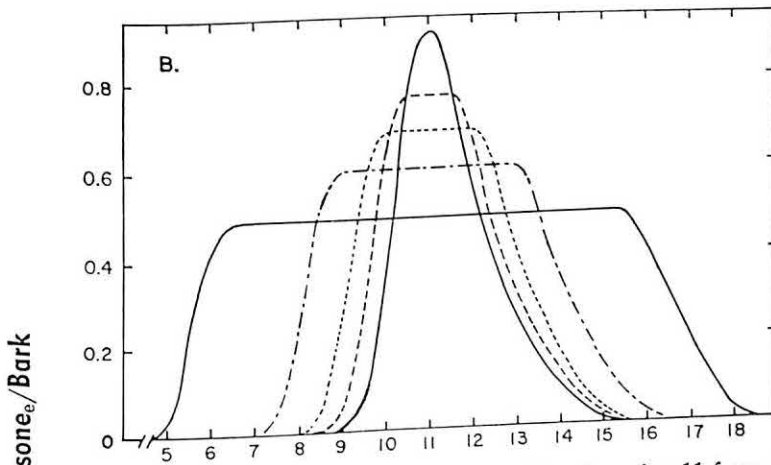
At low values of L_M and L_t , the curves in Figure 7 resemble those in Figure 6. Since, after the correction for the transmission characteristics of the middle ear, L_t is small for all but the low frequencies, perhaps the curves of Figure 7 ought to supercede those of Figure 6. Until it is known, however, which set of curves is more appropriate for the low frequencies, where internal masking is assumed to account for the large absolute threshold, it is expedient to retain the curves of Figure 6, which have already been in circulation for some time. Furthermore, the curves of Figure 7 are for pure tones masked by noise and probably do not accurately describe the growth of the specific loudness of tones masked by other tones or of narrow-band noise masked by tones or by other noises (Zwicker, 1963).

APPLICATION OF THE MODEL TO LOUDNESS SUMMATION AND ANALYSIS

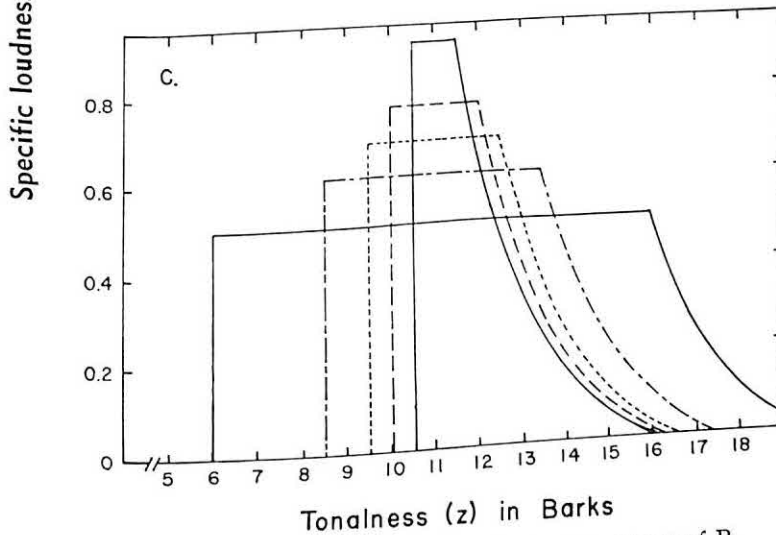
Our model of loudness serves two important and closely related functions. It permits an understanding of the changes in loudness under a variety of conditions involving both summation



8a. Excitation patterns for 5 bands of noise varying in width from 1 to 10 critical bands.



8b. Specific loudness patterns calculated by Equation 11 from A.



8c. Schematic version of the loudness patterns of B.

FIG. 8. Excitation and loudness patterns for bands of white noise at 50-decibel SPL.

and analysis. A convenient transformation of the input stimulus into excitation and loudness patterns provides a means for inferring the interactions within the auditory system. In terms of these interactions, most of the facts of loudness summation may be explained. At the same time the model permits the precise calculation of loudness so that manipulations of the model to predict (or postdict) the results of subjective measurements may be put to rigorous test. (The model is also presented elsewhere [Zwicker, 1959] in a more schematized version for third-octave band filters that simplifies the calculation of loudness where a numerical value is the only requirement as in audio engineering.)

In this section, the model is applied to three types of sound. In each case an important aspect of loudness summation or analysis is explained in terms of the model, and calculated values for loudness are compared to measured values. The first sound is a band of white noise. The second sound is a pure tone partially masked by a narrow band of noise. The third sound is a 4-tone complex partially masked by a wide-band noise.

Loudness of Bands of Noise

The loudness of a band of noise, of constant overall intensity, remains unchanged as its bandwidth is increased up to the critical bandwidth. Beyond the critical band the loudness of the noise usually increases, but at a faster rate for a noise at a moderate SPL than for one at either a high or low SPL. The results of Zwicker, Flottorp, and Stevens (1957) followed this typical pattern, and in the application of the model to their stimuli, the reasons for such results become clear.

Figure 8 presents the excitation patterns for five different bands of noise,

all centered at a geometric mean of 1,480 cps and all with an overall SPL of 50 decibels. These excitation patterns are based upon masking patterns like those in Figure 4. For the noise one critical band wide, the excitation pattern corresponds to that produced by a pure tone or narrow band of noise. For a noise wider than a critical band, the patterns produced by contiguous critical bands have been combined. By means of Equation 11 as graphed in Figure 6, the values L_E are converted over tonalness to the specific loudness N' . These values are plotted as the loudness patterns in Figure 8b and also in Figure 8c. The more schematized versions of Figure 8c facilitate integration and show more clearly how the patterns change with level. The loudness levels computed from Figure 8c were almost identical to those from Figure 8b. Loudness values were also computed for bands of noise at 15 and 100 decibels from the loudness patterns presented in Figures 9 and 10.

It becomes clear from these patterns why loudness is constant within the critical band and why its increase beyond the critical band depends upon level. Consider first Figure 8c. The loudness of each of the 5 noises is proportional to the area under the loudness pattern

$$\left(N = \int_{z=0}^{z=24} N' dz \right).$$

Given the same overall intensity and center frequency, any stimulus no wider than a critical band, even a pure tone, produces a loudness pattern like that shown here for the noise one critical band wide (the highest pattern). Subcritical bands are represented by an invariant loudness pattern because they produce an invariant masking pattern (Zwicker, 1958). Once the critical

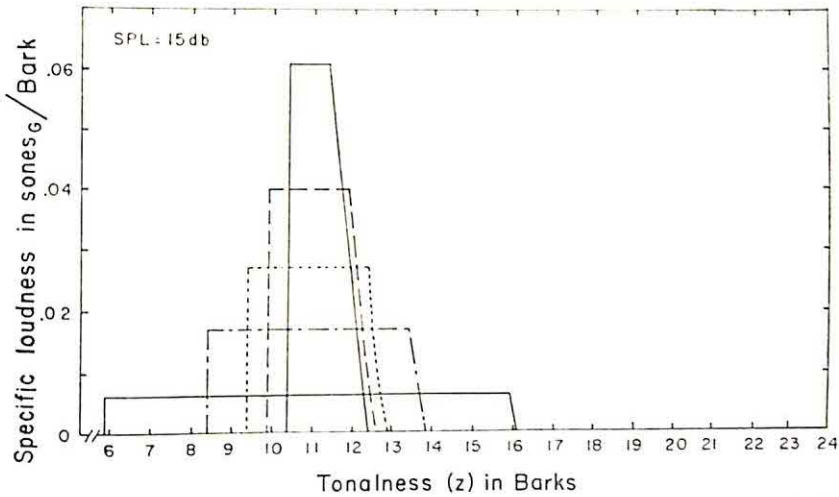


FIG. 9. Specific loudness patterns for 5 bands of noise, all at 15-decibel SPL.

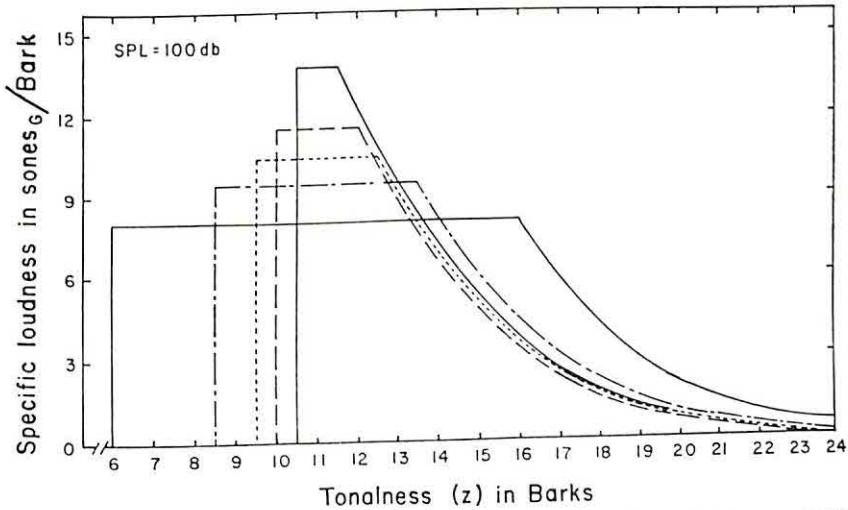


FIG. 10. Specific loudness patterns for 5 bands of noise, all at 100-decibel SPL.

band is exceeded, however, the shape of the loudness pattern begins to change. Its height decreases (owing to the decrease in the intensity of each critical band), and it extends over a greater portion of the z axis. At moderate levels, such as 50 decibels in Figure 8b, the vertical decrease of the curve is more than compensated by the horizontal increase, so that the total area and the total loudness increase. Thus although the excitation and with it the height of the central portions of

the loudness pattern decrease as bandwidth increases, excitation at the lower and upper tonalness values increases sufficiently and enough new elements come into play to raise the overall loudness. Here we see why loudness increases with bandwidth at moderate levels. Why doesn't it increase as much at low and at high levels?

Figure 9 provides an answer for soft noises. Two changes from the 50-decibel noise are evident. First the height of the loudness patterns de-

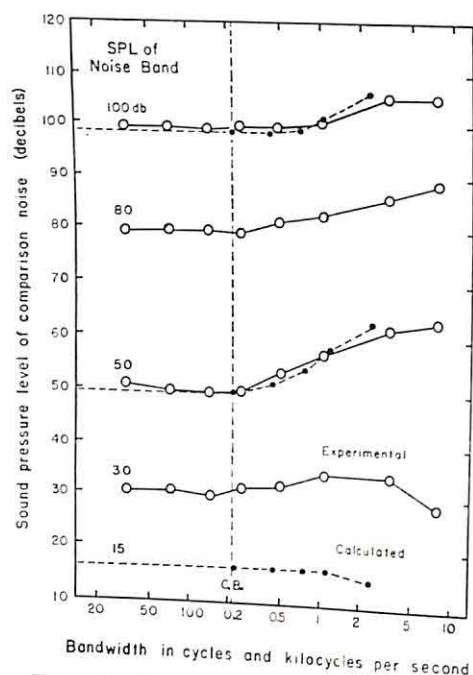


FIG. 11. The dependence of the loudness of a band of noise upon bandwidth. (The parameter on the curves is the overall SPL of the bands of noise. Open circles are values measured in an experiment by Zwicker, Flottorp, and Stevens, 1957. Filled circles are the values calculated from Equation 11. An excellent fit between measured and calculated values obtains.)

creases more rapidly as the stimulus energy is spread out, primarily because the specific loudness decreases with excitation level more rapidly at low than at moderate or high levels, as may be seen in Figure 6. (Close to threshold the loudness of a pure tone changes more rapidly as a function of intensity [Hellman & Zwislocki, 1961; Scharf & J. C. Stevens, 1961].) A second change from the moderate level is that the patterns are narrower and they increase in width somewhat more slowly as the stimulus energy is spread out. Owing primarily to the rapid decrease in the height of the patterns, loudness actually decreases as bandwidth increases beyond the critical band.

At the high levels, the situation is

reversed as a comparison of Figure 10 with Figure 8c reveals. Here, the height of the loudness pattern changes with increasing bandwidth in approximately the same proportion as at the moderate level, but the loudness patterns are spread over considerably larger portions of the tonalness scale at the higher end. The spread is initially so extended that increasing the bandwidth beyond one critical band increases at first only the area at the lower tonalness values. At the same time, the specific loudness is reduced at the middle tonalness values, offsetting any gain from the greater spread. Consequently the loudness does not begin to increase until the stimulating bandwidth extends over 3 to 5 critical bands.

Integration of the areas under the loudness patterns of Figures 8c, 9, and 10 yields values for the loudness level of the bands of noise. These values are plotted as the filled circles in Figure 11. The open circles are the medians of 12 judgments by 12 subjects who adjusted the intensity of a band of noise 210 cps wide to match the loudness of each of the bandwidths shown on the abscissa (data from Zwicker, Flottorp, & Stevens, 1957). The overall SPL of the bands of noise was held constant at 30, 50, 80, or 100 decibels. The calculated and measured values lie close to one another, showing the same invariance of loudness summation within the critical band and the same dependence on level beyond the critical band.

Loudness of Partially Masked Tones

Only with certain types of complex sounds, such as bands of white noise, do all the components fuse to yield an overall loudness. The sound of a plane passing overhead does not fuse with the sound of music coming from the

nearby radio. Yet the presence of the one sound may distinctly affect the loudness of the other. Such a situation arises when a tone and a band of noise are presented together, and the model may be used to understand some of

the interactions that take place within the auditory system and also to calculate the loudness of the tone.

Measurements are available of the loudness of pure tones partially masked by a narrow band of noise (Scharf,

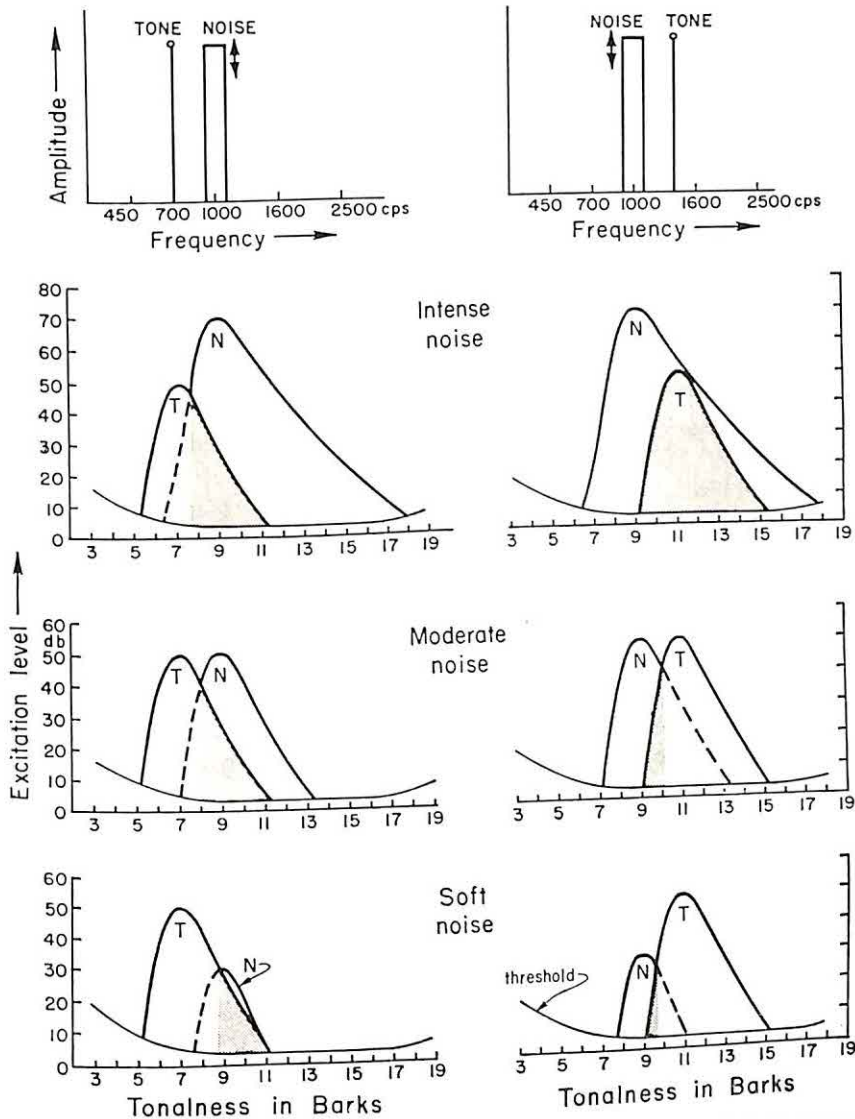


FIG. 12. Excitation patterns for a narrow-band masking noise and two tones. (The idealized spectra for the noise and tones are also shown. In the overlapping [shaded] areas the excitation produced by the tone is suppressed by the noise. Clearly as the noise intensity is reduced, interference with the higher frequency tone decreases more rapidly than that with the lower frequency tone. Thus although the higher frequency tone is completely masked at a noise level where the lower frequency tone is not, once its masked threshold is exceeded, the higher frequency tone grows in loudness more rapidly [Scharf, 1964; adapted with permission of *Acustica*].)

1964). The tones were located at frequencies above and below the frequency range of the noise, which was one critical band wide centered at 980 cps. Whereas complete masking was shown, as usual, to be more effective toward the higher frequencies, partial masking was found to be generally more effective toward the lower frequencies. This change appears paradoxical until the model is used to analyze the excitation patterns produced by the tone and band of noise.

Figure 12 shows the standard excitation patterns for two of the tones and for the band of noise used in the experiments. At the top of the figure are the idealized spectra of the three sounds. Analysis of these patterns is simplified if we assume that at those points where the excitation patterns of the tone and noise overlap, whichever pattern has the higher excitation level completely suppresses the other. (A difference of 3 decibels would actually be required for complete suppression.) In the figure an intense noise completely masks the higher frequency tone, but only partially masks the lower frequency tone. When the noise intensity is reduced, more

of the excitation pattern of the lower frequency tone remains suppressed than of the higher frequency tone. The size of the suppressed area indicates the degree of loudness reduction, i.e., partial masking, so that it is clear that the moderate and soft noises partially mask the lower frequency more than the higher. This difference arises from the asymmetry of the excitation patterns toward the higher frequencies as examination of the patterns reveals.

The model can also be used to calculate the loudness of the partially masked tones. Figure 13 shows the excitation patterns for the masking noise at various levels and for one of the tones. These patterns differ somewhat from the standard excitation patterns; they were taken from the masked audiograms produced by the bands of noise used in the experiment. The same subjects made both the threshold and the loudness judgments. The loudness patterns for the 830-cps tone set at 65-decibel SPL and presented against five different levels of the narrow-band noise are shown in Figure 14. These loudness patterns were constructed by using the curves for the specific loud-

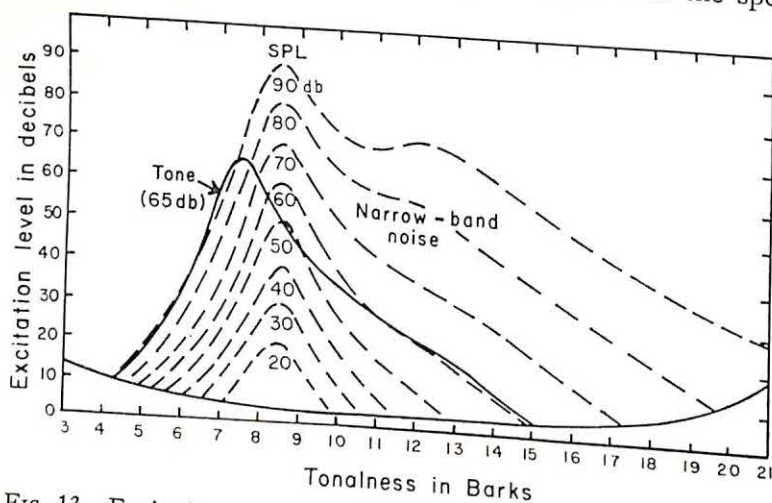


FIG. 13. Excitation patterns for an 830-cps tone (solid curve) and a narrow-band noise at various SPLs.

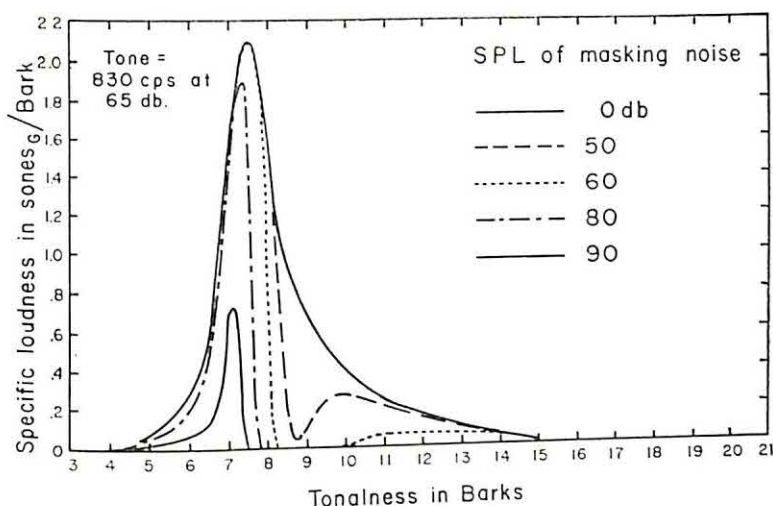


FIG. 14. Loudness patterns for an 830-cps tone (at 65-decibel SPL), partially masked by a narrow-band noise. (The noise was centered at 980 cps and set at the SPLs shown. The loudness patterns for the tone do not change until the noise SPL exceeds 35 decibels. When the noise SPL reaches 96 decibels, the tone is completely masked and is inaudible.)

ness under partial masking (Fig. 7) to convert the pure tone excitation levels of Figure 13 to the specific loudness in $\text{sones}_G/\text{Bark}$. The excitation level L_E of the noise pattern provided the values for L_M required in Figure 7. Loudness patterns (not reproduced here) were also constructed for the 830-cps tone at 45 decibels and also for tones at 980 and 1,355 cps at 45 and 65 decibels. Values calculated from the loudness patterns are plotted for the tones at 65 decibels in Figure 15 along with the measured loudness levels (Scharf, 1964) and the interquartile ranges. (The plot for the tones at 45 decibels is similar to that for tones at 65 decibels and is therefore not reproduced here.) The agreement of the observed with the calculated values is best for the 980-cps tone where the tone and noise were centered on the same frequency so that inaccuracies in the excitation patterns would tend to cancel out. The agreement is good for the 830-cps tone. The agreement is not as good for the 1,355-cps tone, although the general trend of the data is reproduced. Since this tone

was at a higher frequency than the noise, the excitation levels at the higher tonalness values where the noise's excitation pattern had little effect on the tone's pattern were especially important in the loudness calculations. But it is precisely at the higher tonalness values on any given excitation pattern that the levels are most uncertain, for the masking patterns upon which they are based are most variable there both within and among subjects.

Given the difficulty of the loudness judgments and the many steps in the loudness calculation in this complex situation, the model fares well not only in a qualitative analysis (Fig. 12), but also in the quantitative calculations. Perhaps, better calculations could be obtained with further refinements of the model. Nevertheless, using the model as it stands, let us turn to a still more complex stimulus arrangement.

Loudness of a Partially Masked Complex Sound

In the same way that the loudness of a pure tone can be distinguished

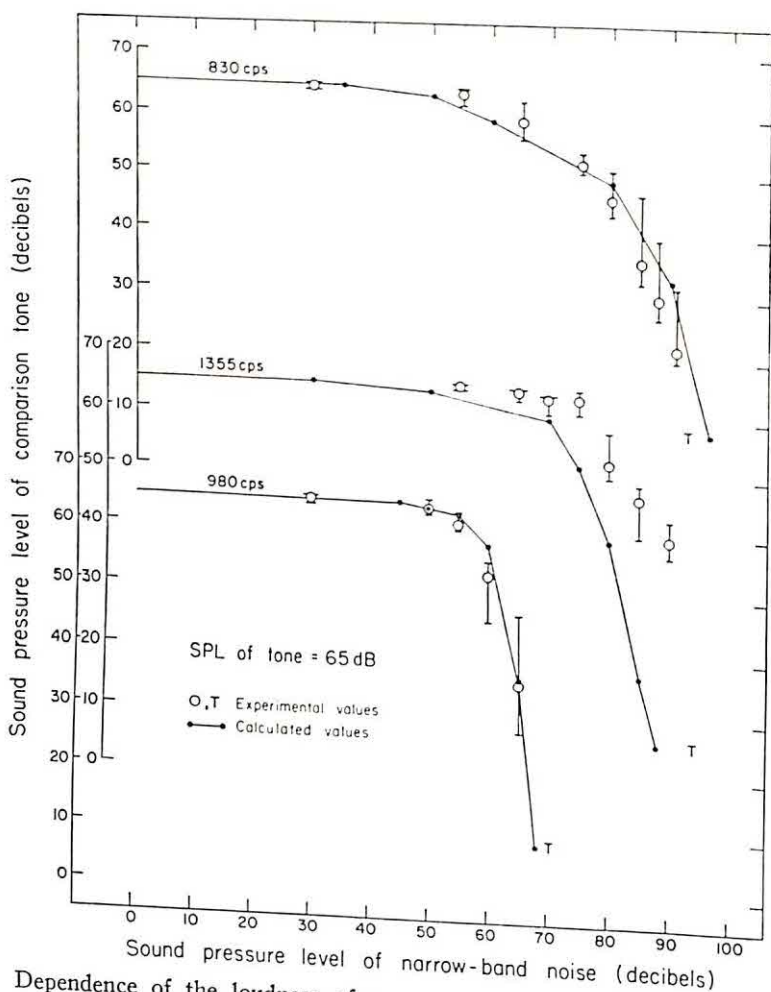


FIG. 15. Dependence of the loudness of a pure tone at 65 decibels on the SPL of a partially masking narrow-band noise. (The ordinate gives the SPL of the equally loud comparison tone which had the same frequency as the masked tone, but was presented in the quiet. The frequency of the tone is the parameter on the curves. Each open circle is the median of 8 to 10 loudness matches by four to five subjects. Interquartile ranges are also shown. The symbol T indicates the measured noise SPL required to mask the tone completely. The filled circles were calculated by means of the curves in Figure 7 from excitation patterns like those in Figure 14.)

from a neighboring band of noise, a multitone complex can be distinguished from a background noise. The loudness of a 4-tone complex has been measured against various levels of a uniform masking noise (Scharf, 1961a). Results showed loudness remains invariant within the critical band. Beyond the critical band, the changes in loudness depend primarily on the sensation level (number of decibels above

threshold) of the complex. Even a moderately intense complex, which in the quiet shows the greatest amount of loudness summation, decreased in loudness as bandwidth increased whenever an intense masking noise was simultaneously presented. How does our model account for this finding?

The excitation patterns for the complex are assumed to remain unchanged in the presence of the noise. Figure

16 shows the excitation pattern for some of the complexes used in the experiment and for a uniform masking noise at an overall level of 55 decibels, which corresponds to an SPL of about 40 decibels in each critical band. These patterns are combined from the standard excitation patterns measured on narrow bands of noise. The individual excitation patterns evoked by the four tones that comprise the complex stand out rather clearly when the tones are spread out as in the 2,200-cps wide complex.

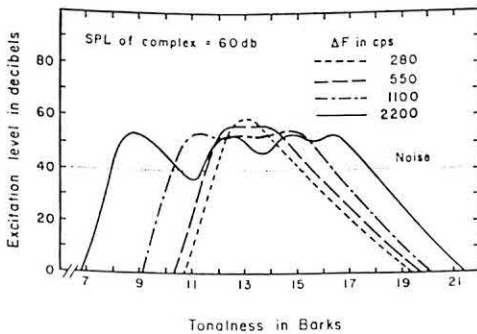


FIG. 16. Excitation patterns for a 4-tone complex at an overall SPL of 60 decibels and centered at 2,000 cps. (Patterns for four different values of ΔF are shown. The dotted line is the excitation pattern for the uniform masking noise when the SPL in each critical band was 40 decibels.)

Whereas the excitation patterns are unaffected by the noise, the loudness patterns may be greatly affected as shown by Figures 17, 18, and 19. Each figure presents the loudness patterns for the complexes in a different level of noise. In Figure 17, where the noise was inaudible at 15 decibels, the loudness patterns grow larger as the complex becomes wider. Noise at 35 or 45 decibels did not produce a measurable change in the loudness patterns for this 60-decibel complex, but with the overall noise level raised to 55 decibels, the patterns do not increase as much with ΔF ; they increase still less

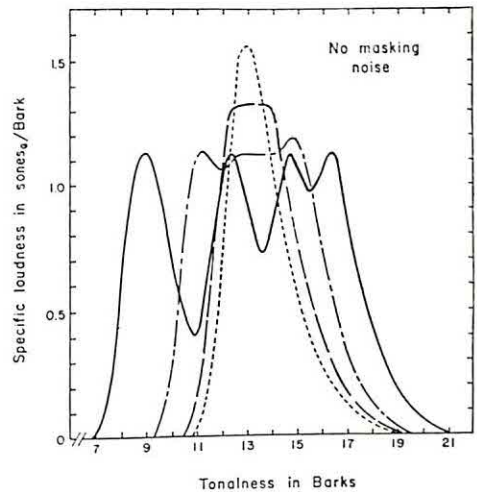


FIG. 17. Loudness patterns for four different 4-tone complexes at 60-decibel SPL. (No background noise.)

when the noise level was 65 decibels (Fig. 18). With the noise at 70 decibels and the complex almost completely masked (Fig. 19), the size of the loudness patterns, and therefore the loudness, actually decreases as ΔF increases.

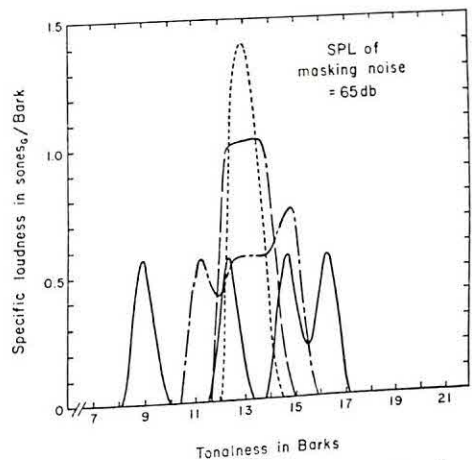


FIG. 18. Loudness patterns for the four complexes presented against a uniform masking noise with an SPL of 50 decibels in each critical band (overall SPL approximately 65 decibels). (Both the height and spread of the patterns are somewhat reduced by the addition of the noise background.)

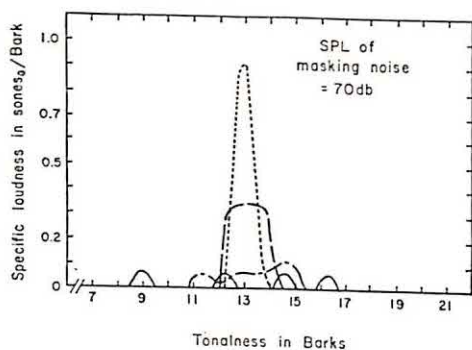


FIG. 19. Loudness patterns for the four complexes presented against a uniform masking noise with an SPL of 55 decibels per critical band (overall SPL approximately 70 decibels).

Loudness diminishes with increased ΔF in the presence of intense noise because near the masked threshold, even more so than near the absolute

threshold, the specific loudness changes very rapidly with the excitation level. Consequently, as the excitation spreads beyond a single Bark, even a small decrease in the level of excitation within each Bark produces a large drop in the loudness pattern. A glance at Figure 7, which is used to transform the excitation levels L_E to the specific loudness values N'_t , shows that near the masked threshold (approximately where the curves cross the abscissa) a reduction of only a few decibels in the excitation level over a single Bark causes a large reduction in specific loudness. Loudness summation appears then to be poorer in the presence of a noise because the noise steepens the loudness functions for the component critical bands of the complex, and

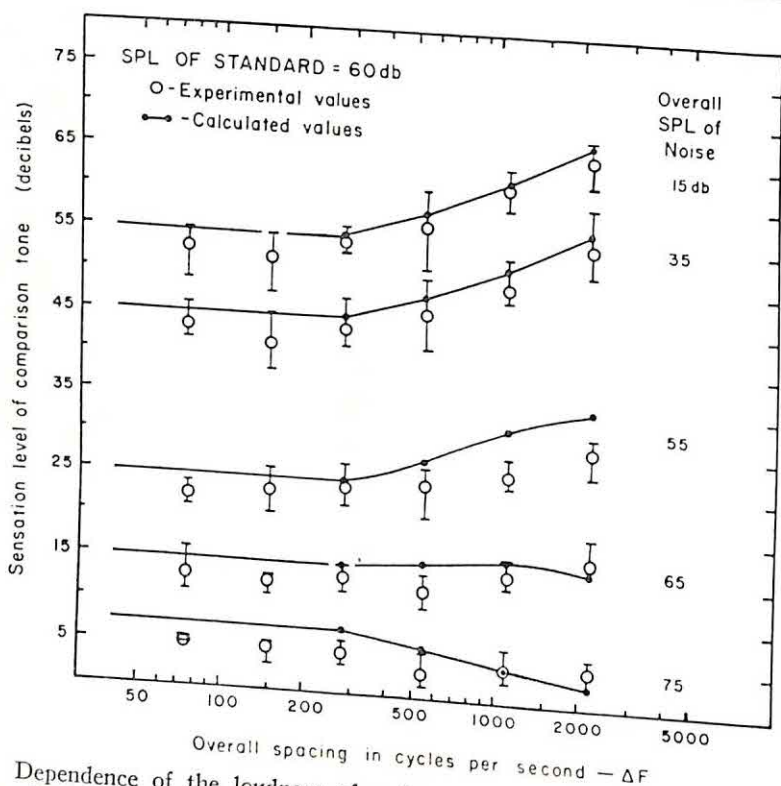


FIG. 20. Dependence of the loudness of a 4-tone complex on ΔF . (The SPL of the complex was held constant at 60 decibels and the SPL of a noise background was varied from 15 to 75 decibels. Open circles are the medians of 12 loudness matches by six subjects; interquartile ranges are also shown. Filled circles are the values calculated from the loudness patterns shown in Figures 17 to 19.)

the more intense the noise relative to the tone, the steeper these functions become.

Now the question remains of how well the calculated loudness values agree with the experimentally measured values. A direct comparison can not be made because the loudness levels of the complexes were not measured in the original experiment. Results were given in terms of the sensation level of an equally loud pure tone heard against the same noise background as the complex. Conversion of the calculated measures to the sensation level of the comparison tone in noise can be made with the help of data recently published on the loudness relations between a pure tone in noise and the same tone in the quiet (Zwicker, 1963). Figure 20 shows the sensation level of the comparison tone that sounded as loud as the complex whose overall spacing is given on the abscissa. The open circles are the medians of 12 judgments by six subjects; the solid lines are fitted to the calculated loudness values. In terms of the absolute values, the agreement is not too good; the calculated values are generally a little too high. However, both the calculated and measured values show the same progressive decrease in the amount of loudness summation as the masking noise becomes more intense, until close to the masked threshold, both sets of values show loudness decreasing as ΔF increases.

Summary of Calculations

With respect to the three types of sounds examined, the agreement between the observed and the calculated loudness values must be considered remarkably good. For each type of sound, the shape of the calculated functions reproduces quite closely that of

the measured functions. Although the absolute values may occasionally differ by many decibels, they are usually in good agreement, especially if we consider that the loudness judgments themselves were highly variable with the interquartile ranges for the various experiments averaging over 4 decibels. It seems unlikely, therefore, that any major modification of the model could eliminate the present discrepancies. As it now stands, the model provides a valid representation of the kinds of physiological processes which can account for most of the psychophysical data on loudness and loudness summation and which also agree with current theories of hearing (cf. Békésy, 1963).

SOME FINAL CONSIDERATIONS

Our model does not try to deal with attentional, attitudinal, or other central factors that certainly affect the experience of loudness. It is intended to apply to normal, unidirectional hearing under conditions of applied attention. Whether, under conditions of attention (such as sleep) that greatly alter the subjective experience of loudness, the model remains a valid representation of neural events occurring at some early stage in the nervous system, we do not know. The whole problem of the role of efferent innervation in the auditory mechanism is under intensive study (Davis, 1962); perhaps some clarification will soon be available. However, even if conditions in the cochlea turn out to be relatively constant, we still do not know whether the summations and interactions pictured in our model are complete at the cochlear level. Indeed our ignorance may be even greater. Ades (1959) writes that, "for the moment, we can only exercise caution in theorizing about the (neural) mediation of loud-

ness . . . [p. 609]." Nevertheless "subjective loudness" is "usually attributed to the number of nerve impulses per second traversing the auditory nerve [Davis, 1959, p. 583]." While our model does not depend on this particular conceptualization (or any physiological theories or data for that matter), the specific loudness is perhaps most easily interpreted on a physiological level as dependent upon the number of neural impulses. Such a view served in the formulation of the model (Zwicker, 1958).

Another underlying assumption that facilitates interpretation of the model is that loudness summation is primarily peripheral. This assumption receives compelling support from Niese's (1960) demonstration that loudness summation takes place monaurally. Niese showed that the loudness of two narrow bands of noise increases as their frequency separation increases beyond the critical bandwidth only if both bands are presented to the same ear. If the two noises are presented separately to each ear, no increase in loudness with increasing frequency separation can be detected. Perhaps more direct evidence for the peripheral nature of loudness summation can be obtained from studies of humans and animals with lesions at various levels within the auditory nervous system. Additionally, studies of interaural effects in multidirectional hearing may provide interesting data.

Physiologically the critical band is more of an enigma than loudness, for we do not know where it is formed or how.² That it is formed upon stimu-

lation and is not already there as some fixed structural unit, something like a variable band-pass filter (cf. De Boer, 1962), is clear. The critical band is a continuous function of frequency, not a step function—it may be measured at all the audible frequencies. Thus the critical band seems to form around the stimulating frequency as stimulation occurs. The formation of the critical band during stimulation suggests the necessity for a period of time during which the locus of stimulation is made known to the auditory system. Just such a temporal lag has been demonstrated by Scholl (1962) in measurements of the masked threshold for short pulses. The usual critical band does not appear until the duration of the pulse exceeds 10 milliseconds. Below 10 milliseconds the critical band appears to become wider and below 5 milliseconds matches the width of almost the whole auditory spectrum. Perhaps the most likely process by which the critical band is formed is of an inhibitory nature, similar to that postulated by Békésy (1960) for the neural unit on the skin and in the eye. If the critical band is a measure of inhibitory processes, we may wonder why the masking pattern measured for a pure tone shows no trace of a discontinuity at the critical bandwidth. Masking (and supposedly excitation) spread out continuously toward the higher frequencies.

The problems raised here are primarily physiological. Perhaps our model can serve to guide research in this area as well as in the psychoacoustical area. Whatever its advantages as a generator of new research, the model provides a comprehensive summary of

any case no one has reported that the average critical band in loudness summation experiments varies in any consistent fashion.

² Indeed, detection experiments using the method of forced choice have led some investigators (Swets, 1963) to infer a critical band whose width depends upon the experimental instructions and the subject's attitude. Such an inference is open to question, and in

available psychoacoustical data on loudness and permits manipulations and calculations that lead to fruitful new insights into the nature of the auditory processes.

REFERENCES

- ADES, H. W. Central auditory mechanisms. In J. Field, H. W. Magoun, & V. E. Hall (Eds.), *Handbook of physiology: Section 1, Vol. 1. Neurophysiology*. Baltimore: Waverly Press, 1959.
- BÉKÉSY, G. v. Neural inhibitory units of the eye and skin. Quantitative description of contrast phenomena. *Journal of the Optical Society of America*, 1960, 50, 1060-1070.
- BÉKÉSY, G. v. Hearing theories and complex sounds. *Journal of the Acoustical Society of America*, 1963, 35, 588-601.
- DAVIS, H. Excitation of auditory receptors. In J. Field, H. W. Magoun, & V. E. Hall (Eds.), *Handbook of physiology: Section 1, Vol. 1. Neurophysiology*. Baltimore: Waverly Press, 1959.
- DAVIS, H. Advances in the neurophysiology and neuroanatomy of the cochlea. *Journal of the Acoustical Society of America*, 1962, 34, 1377-1385.
- DE BOER, E. Note on the critical bandwidth. *Journal of the Acoustical Society of America*, 1962, 34, 985-986.
- EHMER, R. H. Masking by tones vs. noise bands. *Journal of the Acoustical Society of America*, 1959, 31, 1253-1256.
- FELDTKELLER, R. Ueber die Zerlegung des Schallspektrums in Frequenzgruppen durch das Gehör. *Elektronische Rundschau*, 1955, 9, 387-389.
- FLETCHER, H., & MUNSON, W. A. Relation between loudness and masking. *Journal of the Acoustical Society of America*, 1937, 9, 1-10.
- GREENWOOD, D. D. Auditory masking and the critical band. *Journal of the Acoustical Society of America*, 1961, 33, 484-502.
- HARRIS, C. M. Residual masking at low frequencies. *Journal of the Acoustical Society of America*, 1959, 31, 1110-1115.
- HELLMAN, R. P., & ZWISLOCKI, J. Some factors affecting the estimation of loudness. *Journal of the Acoustical Society of America*, 1961, 33, 687-694.
- HOWES, D. H. The loudness of multicomponent tones. *American Journal of Psychology*, 1950, 63, 1-30.
- MUNSON, W. A., & GARDNER, M. B. Loudness patterns—A new approach. *Journal of the Acoustical Society of America*, 1950, 22, 177-190.
- NIESE, H. Subjektive Messung der Lautstärke von Bandpassräschen. *Hochfrequenztechnik und Elektroakustik*, 1960, 68(1), 202-217.
- OETINGER, R., & HAUSER, H. Ein elektrischer Kettenleiter zur Untersuchung der mechanischen Schwingungsvorgänge im Innenohr. *Acustica*, 1961, 11, 161-177.
- PORT, E. Ueber die Lautstärke einzelner kurzer Schallimpulse. *Acustica*, 1963, 13(Beiheft 1), 212-223. (a)
- PORT, E. Zur Lautstärkeempfindung und Lautstärkemessung von pulsierenden Geräuschen. *Acustica*, 1963, 13(Beiheft 1), 224-233. (b)
- ROBINSON, D. W. The subjective loudness scale. *Acustica*, 1957, 7, 217-233.
- SCHARF, B. Critical bands and the loudness of complex sounds near threshold. *Journal of the Acoustical Society of America*, 1959, 31, 365-370. (a)
- SCHARF, B. Loudness of complex sounds as a function of the number of components. *Journal of the Acoustical Society of America*, 1959, 31, 783-785. (b)
- SCHARF, B. Loudness summation under masking. *Journal of the Acoustical Society of America*, 1961, 33, 503-511. (a)
- SCHARF, B. Complex sounds and critical bands. *Psychological Bulletin*, 1961, 58, 205-217. (b)
- SCHARF, B. Loudness summation and spectrum shape. *Journal of the Acoustical Society of America*, 1962, 34, 228-233. (a)
- SCHARF, B. Loudness summation in conductive deafness. Paper H53 in *Proceedings of the Fourth International Congress of Acoustics*. Copenhagen: Harland & Toksvig, 1962. (b)
- SCHARF, B. Partial masking. *Acustica*, 1964, 14, 16-23.
- SCHARF, B., & STEVENS, J. C. The form of the loudness function near threshold. *Proceedings of the Third International Congress of Acoustics*, Vol. 1. Amsterdam: Elsevier, 1961.
- SCHOLL, H. Das dynamische Verhalten des Gehörs bei der Unterteilung des Schallspektrums in Frequenzgruppen. *Acustica*, 1962, 12, 101-107.
- STEINBERG, J. C., & GARDNER, M. B. The dependence of hearing impairment on

- sound intensity. *Journal of the Acoustical Society of America*, 1937, 9, 11-23.
- STEVENS, S. S. Calculation of the loudness of complex noise. *Journal of the Acoustical Society of America*, 1956, 28, 807-832.
- STEVENS, S. S. On the psychophysical law. *Psychological Review*, 1957, 64, 153-181.
- SWETS, J. A. Central factors in auditory frequency selectivity. *Psychological Bulletin*, 1963, 60, 429-440.
- ZWICKER, E. Die elementaren Grundlagen zur Bestimmung der Informationskapazität des Gehörs. *Acustica*, 1956, 6, 365-380.
- ZWICKER, E. Ueber psychologische und methodische Grundlagen der Lautheit. *Acustica*, 1958, 8(Beiheft 1), 237-258.
- ZWICKER, E. Ein graphisches Verfahren zur Bestimmung der Lautstärke und der Lautheit aus dem Terzpegeldiagramm. *Frequenz*, 1959, 13, 234-242.
- ZWICKER, E. Subdivision of the audible frequency range into critical bands (Frequenzgruppen). *Journal of the Acoustical Society of America*, 1961, 33, 248.
- ZWICKER, E. Ueber die Lautheit von ungedrosselten und gedrosselten Schallen. *Acustica*, 1963, 13(Beiheft 1), 194-211.
- ZWICKER, E., & FELDTKELLER, R. Ueber die Lautstärke von gleichförmigen Geräuschen. *Acustica*, 1955, 5, 303-316.
- ZWICKER, E., FLOTTORP, G., & STEVENS, S. S. Critical band width in loudness summation. *Journal of the Acoustical Society of America*, 1957, 29, 548-557.

(Received February 4, 1964)

CONTEMPORARY INSTINCT THEORY AND THE FIXED ACTION PATTERN¹

HOWARD MOLTZ²

Brooklyn College

The observable properties of the Fixed Action Pattern (FAP) as well as the processes assumed by the contemporary instinct theorist to underlie the structure and organization of this allegedly unique response class are described. An attempt is then made to assess the significance of the FAP for psychology. Finally, a critical evaluation of the FAP, in the light of a review of relevant empirical evidence, is presented. On the basis of this evaluation it is concluded that the empirical properties of the FAP do not require the assumptions of genetic encoding and central itemization—assumptions the instinct theorist considers indispensable—but instead can be interpreted in a manner broadly consonant with an epigenetic approach.

Despite the many vicissitudes which the concept of instinct has undergone during the course of its long and multifaceted history, its essential approach to the analysis of behavior has remained virtually unchanged: the postulation of some form of invariant genotypically fixed response nucleus around which the more flexible and plastic components of a behavior repertoire allegedly develop. Holding forth the seemingly attractive promise that such a nucleus would constitute for understanding the structure and organization of a species' repertoire, it is not surprising that the concept of instinct has managed to remain viable

despite repeated attempts to exorcise it from the body of science. The purpose of the present paper is to discuss and critically evaluate its most recent bid for redemption—a bid embodied in the form of a new theory of instinct.

This theory, formulated by several European ethologists of whom the most notable are Lorenz and Tinbergen, has as its core the concept of the Instinctive Movement or the Fixed Action Pattern (FAP). Based primarily, although by no means exclusively, on naturalistic observations of free-ranging species, the concept of the FAP has come to occupy a position of such central importance in contemporary instinct theory that one can hardly discuss the current state of that theory without describing the observable characteristics of FAPs as well as the neurophysiological events which have been assumed to underlie their structure and organization. Accordingly, our first task shall be to delineate briefly those empirical and theoretical properties which have been assumed to render the FAP a distinct class of response events. (Detailed accounts and illustrations of a wide variety of FAPs are to be found in

¹ The present paper was written while the author was in receipt of Research Grant M3855 from the National Institutes of Health, United States Public Health Service.

² Anyone familiar with the writings of Schneirla will recognize the profound and pervasive influence they have had on my thinking about problems of behavioral development in general and about the genesis of species-typic responses in particular. It is a pleasure to acknowledge what is indeed a deep indebtedness. Equally worthy of mention are the critical and penetrating insights offered by Evelyn Raskin—insights from which the present paper profited immeasurably.

Eibl-Eibesfeldt and Kramer, 1958, in Hinde, 1959, and in Moltz, 1962.) We shall then attempt, in the light of a review of the relevant empirical evidence, to assess the significance of the FAP as well as to evaluate the theory into which it has been incorporated. This evaluation will draw upon such valuable appraisals of instinct theory as those of Hebb (1953), Hinde (1954, 1959), Lehrman (1953, 1956), and Schneirla (1956, 1957).

Before beginning, however, two points should be made explicit. First, because the contemporary instinct theorist often refers to himself as an "ethologist" and because he often speaks of ethology as having derived from the study of the FAP, there has been a tendency on the part of American and European psychologists alike to regard all ethologists as instinct theorists. Nothing indeed could be further from the truth. Many investigators call themselves "ethologists" simply to indicate that they are zoologists engaged in the comparative study of animal behavior; they would strongly object to being classified as instinct theorists and sorely disappointed if their contributions came to be regarded as relevant only to instinct theory. For a detailed review of the general "orienting attitudes" of ethology and for a discussion of its history in mediating research, see Hinde (1959).

Secondly, the impression should not be gained that all those who have contributed to the nexus of ideas and concepts referred to herein as "contemporary instinct theory" have remained fixed in their beliefs. There have been apostasies to be sure, most notably that of Tinbergen himself (1955; see also Hinde & Tinbergen, 1958) who has repudiated some of his earlier theoretical formulations. That no mention is made of this in what follows is

due simply to the fact that our task is not to examine the views of any individual theorist nor to discuss the way in which his views might have changed, but rather to evaluate a particular approach to the understanding of species-typic behavior—an approach which continues to have a profound influence on the current thinking of those concerned with developmental problems.

EMPIRICAL PROPERTIES OF THE FAP

Although there is widespread agreement that FAPs comprise a unique class of responses, there is little agreement among ethologists in general or even among instinct theorists themselves concerning the criteria requisite for class inclusion. However, each of the four properties which we shall discuss below in attempting to delineate the empirical character of the FAP has been considered a distinguishing and primary feature and each has in turn posed rather important theoretical issues.

Stereotypy of the FAP

A reasonably precise distinction is often made by ethologists and instinct theorists alike between appetitive behavior and consummatory acts. "Appetitive behavior" is the term used to designate the labile initial components of a behavior sequence, while "consummatory act" is the term used to designate the rigid and terminal components (Thorpe, 1954, 1956).

Those movements or movement patterns designated as FAPs are classified as consummatory acts since they usually comprise the terminal aspects of a response sequence and since they are rigidly stereotyped and constant in form. Their most salient characteristic seems to be their stereotypy; indeed, in this respect, they have been compared with morphological traits,

for, like such traits, FAPs are expressed in a highly invariant manner within a taxonomic group (Lorenz, 1956).

As an example of this invariance we may consider an aspect of the reproductive behavior of the three-spined stickleback. Aeration of the stickleback nest, which is apparently necessary to maintain the oxygen concentration of the surrounding water at a level appropriate for egg development, is accomplished through a movement called "fanning." This consists of using the pectoral fins to exert a forward pressure on the water and simultaneously using the tail and caudal fin to exert a backward pressure equal in intensity to the forward pressure. The sequence of movements performed by these structures is the same on each occasion that fanning is exhibited (Baerends, 1957). Indeed, even under circumstances which apparently force the fish to adopt an atypical position in relation to the substrate, the patterning or coordination of the component response elements remains invariant.

Independence from Immediate External Control

If a movement pattern is to be classified as an FAP, the temporal sequence of muscular contractions comprising the pattern must be independent of afferent regulation. Such independence is evidenced by the fact that the movement will often continue to completion irrespective of changes in external conditions. This property serves to distinguish the FAP from movements termed "taxes" which once elicited continue to be directed by events external to the animal and cease to occur once the external stimulus is removed. The egg-retrieving pattern of the graylag goose, described

in detail by Lorenz and Tinbergen (1938), illustrates the difference between the taxis and the FAP.

A brooding graylag goose, observing that an egg has rolled out of its nest, reacts in a characteristic manner. It slowly rises from the nest and approaches the displaced egg. Upon reaching it, the neck is extended downward and forward so that the undersurface of the bill comes to rest against the far side of the egg. Two distinct movements are then employed which serve to roll the displaced egg in the direction of the nest—a sagittal movement that keeps the egg rolling in the bird's median plane, and a lateral or side-to-side movement that keeps it from deviating too far either to the right or left.

The sagittal movement has been classified as an FAP since its form remains constant despite the irregularities of the terrain over which the egg is rolled and despite differences in the shape of those objects that have been experimentally substituted for the egg. Furthermore, if the egg rolls away from the bill, as occasionally happens in the natural situation, the goose will often continue to perform this movement in the same manner as when the egg was present, indicating that the movement is apparently no longer under extrinsic control.

The lateral movement is classified as a taxis, since it is both evoked and continuously directed by contact of the egg with the undersurface of the bill. Thus, if the egg deviates slightly from the bird's median plane, a compensatory movement either to the right or to the left immediately restores it to its path. If an object that is unlikely to deviate (e.g., a cylinder or light wooden cube) is substituted for the displaced egg, few or no lateral movements are performed; finally, if the egg happens to roll completely free of

the bill, the lateral movement, unlike the sagittal, ceases.

Although the example just given illustrates quite clearly the distinction between the taxis and the FAP, it is difficult in many cases to determine whether a particular movement is directed by external events or is simply released by such events (Baerends & Baerends-van Roon, 1950). Furthermore, the taxis and the FAP frequently occur either simultaneously or in close succession, such interlocking making differentiation additionally difficult (Lorenz, 1937).

At this point it is appropriate to emphasize what has thus far been implicit in our discussion, namely, that the FAP does not possess the characteristics of a chain reaction. Although it may involve a relatively complex pattern of muscle contractions, the FAP cannot be fractionated into successive response links with different external stimulus factors responsible for their evocation. Each component of an FAP, in other words, can be elicited only by the same stimulus or stimulus complex as that which elicits the entire FAP.

Spontaneity

A third property of the FAP is its spontaneity, a term used here to denote fluctuations in threshold that are independent of changes in external conditions. The general rule observed to apply to such fluctuations is that an organism's readiness to perform a particular FAP and the intensity with which that performance occurs are a positive function of the time elapsing since the movement was last evoked.

Van Iersel (1953) demonstrated that when a male stickleback is prevented from fanning for several minutes, a significant increase in the intensity of this activity will occur after the fish is allowed to return to its

nest. This is not due to changes in such extrinsic conditions as the accumulation of carbon dioxide and other gases released by the eggs in the fish's absence, since the same result was obtained when the experiment was repeated with the nest kept completely covered.

Fatigue has also been ruled out as a factor contributing to the fluctuations in the occurrence and intensity of FAPs. If one keeps a cichlid fish in close confinement with conspecifics so that stimuli evoking fighting are continuously present, the thresholds of such agonistic movements as lateral tail beating, spreading the gill membranes, and erecting the median fins are raised (Lorenz, 1956). Neither general fatigue nor fatigue of specific effectors can account for this finding, inasmuch as the fish will at the same time perform other activities involving the same effectors and apparently requiring just as much effort.

The most dramatic bit of evidence regarding the spontaneity of the FAP is its tendency under certain circumstances to occur *in vacuo*. That is, when prevented from occurring for a considerable period by the withdrawal of the external stimulus normally responsible for its "release," an FAP will sometimes be performed in the absence of that stimulus.

For example, van Iersel (1953) reports that fanning is occasionally exhibited in perfect detail by the stickleback in the absence of a nest. Similarly, Lorenz has called attention to the fact that the complex motor pattern involved in the weaver bird's weaving strands for its nest is sometimes performed in the absence of plant fibers that normally serve as the functional object.

Although vacuum activities are performed infrequently and performed under conditions atypical for the spe-

cies, the tendency for less extreme variations in elicibility to occur as a function of the time elapsing since previous performance is as already indicated considered to be characteristic of all FAPs. Moreover, this characteristic is often emphasized in the attempt to distinguish it from reflexes or reflex-like movements (Eibl-Eibesfeldt & Kramer, 1958; Lorenz & Tinbergen, 1938; Thorpe, 1954).

Independence from Individual Learning

A fourth criterion for membership in the class of FAPs is the exclusion of the possibility that the movement or movement pattern has been specifically learned prior to its first occurrence. Lorenz (1956), for example, maintains that even wide fluctuations in environmental conditions during ontogeny will in no way alter the FAP provided the health of the organism is not impaired.

The isolation technique is frequently regarded as the critical method for assessing the influence of learning (Eibl-Eibesfeldt, 1961). In its simplest form, this technique consists of removing an animal at the time of birth or hatching from members of its own species and then determining whether a given response is subsequently performed in a manner identical to that shown by animals reared with conspecifics. Thus, Tinbergen (1942) and Cullen (1960) report that male sticklebacks reared in isolation performed a typical zigzag courting dance before they had ever seen another stickleback; in fact it was performed on the first occasion that they were introduced to a cardboard model of a gravid female.

The appearance of many movements designated as FAPs very early in ontogeny is also considered evidence that they could not have been learned in

any ordinary sense of that term. Thus, the side-to-side head movement of the human infant is present at birth (Prechtl, 1958), and the gaping response of the thrush is present at hatching (Tinbergen & Kuenen, 1939).

Additional evidence that the FAP is dependent neither on individual learning or practice, comes from observations that the FAP occurs even when certain structures involved in the movement are absent. For example, in one species of surface-feeding ducks, courtship involves a preening movement by which the drake exposes and wriggles a brilliantly colored tertiary feather. Lorenz (1955) reports the case of a drake that had, for unknown reasons, failed to develop this feather but which nonetheless persisted in performing the preening movement. Similarly, FAPs can appear in ontogeny long before they possess functional utility and hence long before a (presumably) necessary condition for learning is present. For example, a young gosling in an aggressive encounter holds its adversary in exactly the same position in space and uses exactly the same beating movement as the adult goose despite the fact that its wings are not large enough to touch its opponent (Lorenz, 1956).

This last example reveals the narrow conception of learning held by most instinct theorists. They reason that, if no source of reinforcement can be identified in connection with a particular response, then the response is perforce not learned (Eibl-Eibesfeldt & Kramer, 1958; Lorenz, 1955). But, despite their tendency to equate the absence of reinforcement with the absence of learning, they have admittedly pointed to many aspects of animal behavior which appear to be nonlearned. They have not, however, thereby demonstrated the reality of behavior ele-

ments that are uninfluenced during ontogeny by extrinsic stimulus conditions, or by what may be more simply labeled experience. Subsequently, we shall have occasion to argue that learning and experience are not equivalent agents in the determination of behavior.

THEORETICAL PROPERTIES OF THE FAP

In our delineation of the empirical characteristics of the FAP, it thus far has been treated solely as a descriptive concept. Now, however, we must turn to an analysis of some of the mechanisms or processes which have been inferred by instinct theorists to underlie the development and organization of the FAP and which in turn explain its empirical properties.

Genetic Encoding

Instinct theorists contend that an organism inherits FAPs just as it inherits morphological structures (Eibl-Eibesfeldt & Kramer, 1958). Although this position does not imply that instinct theorists believe that FAPs develop independently of a specific embryogenic environment, it does imply that the elements of the *genetic constitution* participate in or contribute to the FAP in a functionally congruent manner.

Indeed, Lorenz and others (e.g., Thorpe, 1961) have been quite explicit in maintaining that each FAP is *genetically encoded*, and that such encoding is expressed in the organization of neural centers which serve to control and coordinate the sequence of muscle actions involved in performance. Bullock (1961) has stated the instinct theorists' position on this point quite precisely:

It seems at present highly likely that for many complex behavioral actions the nervous system contains not only genetically deter-

mined circuits but also genetically determined physiological properties of their components so that the complete act is represented in coded form and awaits only an adequate trigger, either internal or external [p. 55].

The recent work of Miller (1957), Andersson (1953), and Harris, Michael, and Scott (1958) has, of course, demonstrated the close association between neural loci and specific responses. But instinct theorists go much further, for they speak of the FAP as being not only centrally integrated but centrally blueprinted—a blueprint that is directly provided for in the growth process itself. And they contend that it is by virtue of the intrinsically established blueprint that the FAP—as a temporally patterned system of response elements—can appear without having been ontogenetically organized. Indeed, since such systems are considered to be genetically specified down to the smallest detail, it is understandable that the influence of experiential events in determining their configuration and phenotypic expression is held to be not only unnecessary but ineffective (Lorenz, 1937, 1956). And it is equally understandable, granting once more the specific encoding, that they should speak of FAPs as constituting the nucleus of an animal's response repertoire—a nucleus which, although overlaid with acquired elements during ontogeny, nevertheless retains its distinctive character in the behavior of the adult organism.

Action Specific Energy

As already pointed out, instinct theorists have been particularly impressed with the spontaneity of the FAP, often regarding it as that property which renders the FAP a unique behavioral event. To Lorenz, this spontaneity suggested that each FAP

possesses its own source of energy which accumulates in that locus of the central nervous system responsible for its coordination. This accumulation is assumed to occur while the particular movement is quiescent and to be expended when the movement is discharged. Tinbergen (1951) has also proposed an essentially similar reservoir-model of motivation involving a hierarchically organized system of neural centers responsible for the activation of a chain of functionally related activities (e.g., those involved in reproduction).

Of particular theoretical interest with respect to the reservoir-type model of motivation is the assumption by the instinct theorist that motor impulses are endogenously produced in a specific center and that the basic pattern of production not only occurs independently of afferent inflow but is in fact refractory to modulation by such inflow. In maintaining that the spontaneity of the FAP is a manifestation of such a purely central automaticity, the instinct theorist points to cardiac and respiratory activities, as well as other basic functions, and considers evidence regarding their spontaneity as providing clues for understanding the neurological basis of complex behavioral spontaneity.

In this connection, the studies of Adrian (1931), von Holst (1934), Maynard (1955) and others are frequently cited. Adrian and Buytendijk (1931), for example, concluded that respiratory activity is endogenously generated after finding that potential changes recorded from the isolated brain stem of the goldfish exhibit the same frequency as that of normal breathing movements. Maynard (1955) found that when the factors which influence heart beat in the lobster are held constant, a complex peri-

odic pattern of impulses is still generated by the cardiac ganglion.

It obviously requires a long inductive leap to go from the comparatively simple functioning of an isolated segment of the nervous system to the temporally organized behavior of the intact animal. But it is precisely this leap that the instinct theorist has taken.

SIGNIFICANCE OF THE FAP

Before attempting to analyze the psychological implications of the FAP, two sources of misunderstanding, each of which has led to some confusion concerning the significance of the FAP, must be clarified. In the first place, it should be apparent that the FAP is not to be identified with response patterns which have been traditionally regarded as instinctive and which have usually been given such global designations as maternal behavior, filial behavior, migratory behavior, etc. What must be emphasized is that each of these labels refers to a myriad of functionally related activities of which only a few might be both empirically and conceptually identical to the FAP. To the author's knowledge, no contemporary instinct theorist ever considered maternal behavior, for example, to be innate in the same sense that the FAP is considered to be innate. Certainly, it is irrelevant to criticize either the reality or the significance of the FAP by showing that patterns like maternal behavior, or any pattern of such heterogeneous composition, is experientially organized and consequently could be neither genetically encoded nor endogenously generated.

Secondly, assessment of the FAP must be twofold, since its value as an empirical concept is independent of the validity of those hypothetical mechanisms assumed by instinct theorists to underlie its structure and organiza-

tion. In discussing significance, we must, therefore, clearly distinguish between the empirical and theoretical roles which the FAP plays.

Significance of the FAP as an Empirical Concept

As a class of response events, FAPs appear to be important empirically in at least three respects: taxonomically, evolutionally, and genetically.

Consider first their taxonomic value. A behavior pattern that possesses the phenotypic properties of the FAP would seem ideally suited to serve as a classificatory device. The particular virtues of stereotypy and resistance to ontogenetic modification should make it possible for the FAP to provide taxonomic information both of a diagnostic and of an associative nature (Mayr, 1958).

Consistent with this claim is the fact that the taxonomist no longer relies exclusively on morphological data; indeed, when structural and behavioral evidence lead to conflicting classificatory decisions, he is often inclined to give greater weight to the latter. And, relevantly enough, many of the responses thus employed appear to be phenotypically akin to FAPs, if not phenotypically identical with them. Crane (1941, 1957), for example, used certain display movements to classify species of fiddler crabs—species that are so structurally alike that a hand lens and a trained eye are required to differentiate them anatomically. Similarly, Hinde (1955), Lorenz (1958), and Tinbergen (1959) have used motor patterns involved in aggression and courtship to elucidate relationships among some closely related avian taxa.

Secondly, the FAP can be used as an instrument in studying behavioral evolution. The morphologist interested in reconstructing phylogenetic

histories has many available traits with which to work—traits that are quantitatively delineable, ontogenetically stable, and interspecifically differentiable. The student of behavior who is devoted to the same pursuit is far less fortunate, since response traits suitable for evolutionary analysis are rare. Again the FAP should prove valuable. The readiness with which it can be discriminated from other components of a species' repertoire, the degree to which it resists ontogenetic modification, and the manner in which it is distributed among related taxa (neither too conservatively nor too divergently) make it ideally suited as an object of phylogenetic study.

Reports already in the literature support this evaluation. Hinde and Tinbergen (1958), and Tinbergen (1954), for example, have offered stimulating evolutionary analyses of certain display movements in birds and Baerends and Baerends-van Roon (1950) have been able to carry out a parallel study for cichlid fishes. Similar analyses have also been performed on many elements of invertebrate behavior that appear empirically akin to FAPs: Crane (1941) on the courtship response of fiddler crabs; Evans (1953) on the wing movements of wasps; and Blest (1960) on the settling behavior of moths, to mention only a few.

A third empirical consequence arising from the discovery of the FAP is its utility in the study of gene-behavior relationships. As Fuller and Thompson (1960) have stressed, selection of appropriate phenotypic measures is one of the most pressing and significant problems for the psychologist interested in behavior genetics. Here also FAPs, because of their environmental stability, their stereotypy, their particulate and quantitatively delineable character, provide

highly satisfactory analytic units for psychogenetic research. Several attempts (e.g., Dilger, 1959, 1962; Hinde, 1956; Ramsay, 1961) have already been made to employ them in just this capacity.

Significance of the FAP as a Theoretical Concept

In its role as a theoretical concept, what possible contributions does the FAP make toward increasing our understanding of behavior? To answer this question, let us accept provisionally the theoretical inferences which the instinct theorist has drawn and examine some of their implications.

It will be recalled that FAPs are conceived to constitute the nucleus of an animal's response repertoire—a nucleus that is encoded in the genome and that subsequently interlocks with all acquired elements of behavior. Although the relationship between these nuclear elements and the more complex learned patterns of behavior has not as yet been elaborated for any species, the instinct theorist nonetheless has contended that analysis of a repertoire into component elements—an essential first step in understanding the behavior of any species—must proceed from the genetically given to the individually acquired. It is considered literally impossible to understand learning, for example, without first understanding the innate substructure upon which it is based and by which it is in turn affected (Lorenz, 1960; Tinbergen & Perdeck, 1951).

One implication then of the theoretical inferences that have been drawn from the FAP is a policy for conducting research: the nuclear elements of a response repertoire must first be distinguished, their properties analyzed, and their influence on all other components investigated. After completion of these tasks, ontogenetic re-

search should then be undertaken but only in relation to those response systems designated as acquired. Obviously, there would be no need to study the cumulative effects of organismic-environmental interactions on innate responses; genetically encoded and centrally blueprinted, their organization should not be influenced to any significant extent by experiential events.

Another consequence of the FAP as a theoretical concept derives from its alleged capacity to provide an index of certain neurophysiological events. It will be recalled that each FAP is considered to be centrally preformed in the sense that a genotypically fixed isomorphism is assumed to exist between its phenotypic properties and the anatomical and physiological characteristics of its neural coordinating center. On this basis, instinct theorists have concluded that it should be possible to make relatively detailed inferences about the nature of such centers from observations carried out at the molar level (Lorenz, 1960).

The opportunity to proceed from the molar to the molecular in so direct and specific a fashion would undoubtedly be of inestimable value in studying neural processes. A temporally integrated response pattern, isomorphic with a localized encapsulated brain region, would provide an instrument, hitherto unavailable, for rendering molar properties into physiological mechanisms. Indeed, even if we were convinced that the translation would not be simple, the very assumption of an isomorphic relationship of the kind postulated by the instinct theorist has profound empirical consequences—consequences quite different from those which would result if we were to assume that FAPs, like most other temporally integrated responses, are probably determined by a complex of functionally heterogeneous neural processes

which occur in anatomically diverse brain regions and which do not correspond either topologically or topographically to the behavioral events they underlie. Thus, expecting functional congruency between the molar and molecular levels makes us prone to homologize all response elements having analogous functional characteristics. As Lehrman (1953) has pointed out, this results in an approach that is essentially anticomparative insofar as phenotypic resemblances tend to be abstracted from diverse phyletic levels and translated into identical physiological mechanisms without regard for differences in species capacities.

There is a third theoretical consequence, still more far reaching in its implications for psychology than the two we have discussed thus far. If the FAP is a temporally integrated behavioral entity which is uniquely organized and with characteristic properties specifically derived from this organization, then laws singularly applicable to the FAP would be necessary to accommodate the fact of its existence. Such a prescription obviously results in a search for novel functional relationships, and this is what the instinct theorist has evidently tried to do. On the one hand, he has attempted to show that responses identified as FAPs cannot be subsumed either by the laws of learning or by the laws of reflex action. On the other hand, he has made an effort to study what he considers to be innately determined perceptory mechanisms—mechanisms which, upon being activated by specific environmental stimuli, presumably function to release the FAP (e.g., Lorenz, 1960, 1961; Tinbergen & Perdeck, 1951).

CRITICAL EVALUATION OF THE FAP

Critics of instinct theory do not deny the reality of the FAP, nor do

they belittle the importance of studying the FAP in understanding vertebrate behavior. The questions which have been asked revolve first about the theoretical constructs used to explain the FAP: do the empirical properties of the FAP demand the kind of hypothetical neurophysiological events deduced by the instinct theorist or can these same properties be subsumed under what might be called an epigenetic approach? A corollary to this question is whether these species-typic responses require the formulation of special developmental laws or can they be integrated within a single developmental theory that embraces acquired responses as well?

We have already noted that the empirical properties of the FAP have been interpreted to support conclusions concerning genetic encoding and central itemization. At the outset, it must be emphasized that there is no evidence, either genetic or embryologic, which in principle contraindicates the possible validity of this interpretation. Nor is behavioral evidence available which would establish conclusively that responses designated as FAPs cannot be innate in the very sense in which the instinct theorist has used that term. It is, however, meaningful to inquire whether the empirical features of the FAP can in fact be explained without recourse to propositions involving the idea of genetic preformism.

Since most FAPs have not been subjected to experimental study, they cannot, by and large, be used to exemplify the manner in which ontogenetic processes can function in determining the structure and organization of species-type behavior. Our analysis, therefore, will frequently draw on research involving responses which are only crudely analogous to the FAP. As a consequence, the evidence

and arguments presented cannot rule out the possibility that the instinct theorist may be correct; but it does provide us with a reasonable and perhaps more cogent alternative for conceptualizing the FAP.

Alternative Routes to Stereotypy

Let us begin by trying to understand how a temporally integrated response exhibited in a virtually identical manner by all members of a taxon could arise even if the response itself were not genetically encoded. Of the number of possible alternative sources for such stereotypy, the most obvious lies in the environment itself.

Involved in the biotic province of any species are certain environmental constancies which may function to produce parallel responses in all normal members (Hebb, 1953; Lehrman, 1953). As Lehrman in particular has pointed out, this homogenization can occur either as a consequence of such constancies channeling ontogeny in congruent directions so that only certain developmental possibilities are realized or of their providing each individual with identical avenues for response expression so that, even when alternate forms of the same response are present in the repertoire, only one is likely to be manifest.

Thus Harker (1953) reports that the highly stereotyped diurnal activity rhythm exhibited by the adult mayfly does not occur unless the individual organism has been subjected to at least one 24-hour light-dark cycle prior to the termination of the larval period (see also Brown, 1959). Thorpe and Jones (1937) showed that the recurrent selection of the flour moth by the parasitic ichneumon fly, *Nemeritis canescens*, for the purpose of oviposition is significantly changed in *Nemeritis* offspring when they are reared on a different host. Apparently, host

selection by an adult is in part a consequence of exposure to the specific chemical and nutritional milieu which *Nemeritis* experiences while in a pre-imaginal stage. And finally, Van der Kloot and Williams (1953a, 1953b) have pointed to characteristically invariant environmental factors which influence the cocoon-construction of *Cecropia* silkworms. Of special interest here is the fact that, when the caterpillar of this lepidopteran species is deprived of the usual physical support (an upright twig crotch) on which to spin its finely tapered cocoon, it will use spinning movements which are quite atypical, resulting in a cocoon markedly distinct from the normal.

The anatomical and functional properties of peripheral structures also provide a possible explanation of behavioral stereotypy, for such structures can function directly to render specific response patterns invariant. Perhaps the clearest exemplification of the detailed interrelationships which can exist between response elements and associated peripheral mechanisms is provided by the work of Davis (1957) on the spotted and brown towhees. These birds forage for food by using a scratching movement which is performed in a highly stereotyped but essentially different manner in each species. Scratching by the brown towhee occurs with a backward thrust that is noticeably less powerful than that manifested by the spotted towhee and incorporates a pronounced lateral component that its congeneric counterpart does not exhibit. Davis has argued convincingly that interspecific differences in the expression of the scratching movement are directly determined by characteristic differences in the osteology and myology of the hindlimbs.

In the present context, it is important to emphasize that functionally

similar stereotyped responses need not be governed by the same causal events, as illustrated by two instances of invariant behavior occurring at widely different levels of the phyletic scale. The jellyfish, for example, exhibits certain characteristic feeding movements which appear to be a direct consequence of the functional properties of its medusoid nerve-net system and the spatial arrangements of its tentacles and manubrium (Maier & Schneirla, 1935). The great tit (Paris major) also exhibits certain species-typic feeding movements but, in contrast to the jellyfish, these movements develop over many ontogenetic pathways and involve more complex interrelationships between intrinsic and extrinsic factors.

Thus, as Schneirla has repeatedly emphasized (e.g., 1946, 1956), stereotypy can arise from causes which operate directly and are immediately apparent or from causes which are considerably more subtle and involve the operation of disparate ontogenetic mechanisms. Our discussion of environmental constancies and structural equipment as conceivable sources of stereotypy does not mean, therefore, that one or the other of these factors is involved in all instances of behavioral rigidification or that their mode of influence is always the same. Not only may there be additional modes of origin but it is also impossible to generalize with respect to the manner in which any one of them operates to homogenize different response patterns at different phyletic levels.

Innateness and the Distinction between Learning and Experience

Let us turn now to consider another characteristic of the FAP—independence from individual learning—and attempt first to see what this criterion means and then to determine whether

independence as such necessarily entails the assumption of innateness.

As we have already pointed out, the instinct theorist seems to conceive of learning in a way that is entirely too restrictive. When a conventional reinforcing agent does not seem to be involved, he is likely to assume that learning could not have occurred. In view of the controversial status of the reinforcement issue in learning theory itself, the absence of reward can hardly be considered as *prima facie* evidence for the absence of learning. But even if we broaden the classificatory criteria for distinguishing the learned from the nonlearned and even if we were to grant that these criteria may change as our understanding of the learning process expands (Beach, 1955; Verplanck, 1955), it still seems reasonable to designate as nonlearned those responses whose development cannot be explained in terms of any extant learning paradigm. Viewed in this way, the FAP can be classified—at least provisionally—as a nonlearned response.

The question which remains to be answered is whether genetic encoding becomes the only plausible way to conceptualize the manner in which nonlearned behavioral elements can arise in ontogeny. The answer is clearly in the negative, for we know that experiential interactions not presently included within the scope of learning can influence the structure and organization of behavior in specific and important ways (Schneirla, 1956, 1957; Schneirla & Rosenblatt, 1961).

Schneirla has used the term "experience" to denote a wide range of stimulative involvements which determine trace effects varying in complexity from specific physicochemical changes in somatic tissue to general functional integrations in the central nervous system. As he goes on to

point out, the manifestations of such effects, in turn, can represent very different levels of behavioral organization. Thus, compare host selection in certain arthropods as influenced by the composition of the medium on which the larvae are fed (Thorpe & Jones, 1937) with the influence of early patterned light stimulation on form discrimination and interocular transfer in vertebrates (Chow & Nissen, 1955; Riesen, 1951, 1958, 1960). Again, contrast the reciprocal stimulative associations ("trophallaxis") apparently essential for the organization of certain insect colonies (Schneirla, 1946) with the development of avoidance responses in mammals as effected by isolation from conspecifics (Melzack & Scott, 1957).

Each of the cases just cited entails ontogenetic processes which, currently at least, are not describable as learning but which nonetheless depend on some type of interaction between the organism and its sensory environment. To clarify the distinction involved here, let us consider one example in detail. Recent studies (Hebb, 1946; McBride & Hebb, 1948; Melzack, 1952; Riesen, 1961) of the genesis of emotional behavior in vertebrates indicate that intense fear can be elicited by strange, although innocuous, visual stimuli. What is necessary for this development is commerce with a structured sensory environment—commerce in which the very fact of exposure seems sufficient to establish the visually familiar. It is evident that the animal does not have to *learn* to discriminate the familiar from the unfamiliar nor does it have to undergo avoidance conditioning before it comes to fear stimuli which differ markedly from those already encountered. Once early contact delimits the perceptually typical, presentation of an incongruous stimu-

lus combination is all that is necessary to evoke an emotional response.

Admittedly, the term experience is difficult to define precisely, but some such term is evidently needed to designate instances, such as those just cited, in which involvement with the ontogenetic milieu operates to structure behavior through channels which do not appear to depend on learning as currently conceived.

We have already mentioned the isolation technique as a means for assessing the influence of learning. What about its value in assessing the influence of experience? There is no doubt that, if proper control conditions are employed, the isolation technique can provide evidence as to the importance of social interaction as a developmental variable. When Cullen (1960), for example, showed that male sticklebacks reared in isolation performed the typical zigzag "court-ing" dance on the first occasion that they were presented with a model of a gravid female, she thereby demonstrated that, whatever else is necessary for the ontogenesis of the dance, experience with conspecifics is not. But to say that the zigzag dance does not develop out of the interaction with species members is obviously not equivalent to asserting that the "whole 'blueprint' of a specifically structured receptor and effector apparatus is contained in the genoma [Lorenz, 1961, p. 183]." This unwarranted deductive leap—one which the instinct theorist often makes on the basis of an isolation experiment—results from the failure to realize that an animal removed from the company of species members is not necessarily removed from other nonsocial environmental effects which in themselves might be sufficient for the organization of the observed behavior pattern.

Indeed, with respect to the genesis

of certain response patterns, the isolation and normal situations may not be significantly different in that they may both provide conditions guaranteeing the occurrence of exactly the same critical experiential involvements. As Lehrman (1953) has put it, "The important question is not 'Is the animal isolated?' but *From what* is the animal isolated? [p. 343]." For example, there is compelling evidence (Birch, 1956; Schneirla & Rosenblatt, 1960; Steinberg & Bindra, 1962) that, in the rat and cat, certain self-stimulative processes, particularly those involved in licking the posterior parts of the body and ingesting vaginal fluids, are implicated in the development of maternal behavior. Separation from conspecifics could hardly be expected to inhibit the occurrence of such processes any more than the presence of conspecifics would be expected to facilitate their expression.

We do not wish to impugn the value of the isolation technique nor to deny the importance of determining whether or not social interactions is critical for the organization of species-typic responses. We merely wish to point to the obvious fact that social interaction is not the only channel through which extrinsic stimulative effects can participate in the genesis of nonlearned behavior and consequently that the isolation experiment cannot be considered decisive in relation to the question of whether a particular response element is genetically encoded or experientially organized.

Independence from Immediate Afferent Control and Neurogenesis

It will be recalled that freedom of the FAP from immediate afferent control has been adduced to support two theoretical propositions: that the component response elements of the FAP are centrally coordinated by virtue of

genetically organized neural circuits, and that the threshold fluctuations characteristic of the FAP arise solely from the accumulation and discharge of endogenously generated motivational impulses.

The problem of how the nervous system becomes structured to support the orderly relations among the constituent parts of a movement pattern is not to be confused with the problem of how central coordinative mechanisms, whether maturationally organized or experientially established, participate in energizing such patterns. The two problems are obviously distinct and warrant separate consideration. We shall treat the first of these in the present section.

To begin with, it should be emphasized that no contemporary student of behavior would maintain (contrary to the contention of Lorenz, 1960) that the embryonic growth process simply leaves an equipotential homogeneous network out of which every neuronal linkage must subsequently be forged by training or experience. At the very least, all would agree that the basic architecture of the nervous system is determined by intrinsic developmental forces—in other words, that the location of nuclear centers and the topographical arrangement of fiber tracts become established during embryogeny in species-typic fashion (Sperry, 1951, 1958). But for the instinct theorist, the role inherited structures play in the organization of the FAP is highly specific—there are coordinating structures in the brain and spinal cord containing genetically patterned circuits which determine the FAP down to its smallest detail. For the critics, in contrast, the intrinsic contribution is very much less determinate and, with the possible exception of mechanisms underlying certain simple reflex reactions, does not include neuronal anlagen suf-

ficient in themselves to provide for the expression of any temporally integrated response. Rather, the growth process, in establishing the general structural and functional plan of the nervous system, is conceived to provide a substrate against which neural configurations subserving coordinated behavior gradually become refined and elaborated through environmental interactions—a substrate, in other words, that essentially determines the directions and limits of a response repertoire but not its details.

With this distinction in mind, we can now turn to the theoretical issue raised by evidence of independence from immediate afferent control. When a response is released but not guided by extrinsic stimuli, it is reasonable to assume that an organized central system has been activated and that the properties of this system influence the phenotypic character of the observed behavior. It is quite likely that such systems underlie the FAP, for, in addition to the presumptive evidence supplied by field observations, there are experimental results which support the same conclusion. Thus Hess (1949) and Hess and Brugger (1943) have demonstrated that, in the cat, punctate electrical stimulation of diencephalic loci can evoke discrete behavioral items normally involved in eating, sleeping, and fighting. More recently, von Holst and St. Paul (1960, 1962) have been able to produce a broad range of movements by stimulating a variety of localized points in the brain stem of domestic fowl. The movements elicited were species-typical and functionally identical with those exhibited in caring for the young, in pecking, and in flight from enemies.

The instinct theorist considers data of the kind just cited to indicate the existence of endogeneously predetermined neuronal systems sufficiently

organized to integrate the component elements of the FAP into an orderly spatial and temporal amalgam. However, no really critical evidence has as yet been adduced in support of this possibility.

First, the fact that punctate stimulation can effect the discharge of an integrated response is in itself not decisive with respect to determining the nature of antecedent neurogenic events. Miller (1957, 1958, 1960), for example, was able to elicit a previously learned bar-pressing response in satiated rats by stimulating ventromedial nuclei in the hypothalamus. In this case, it is obvious that extrinsic induction rather than intrinsic self-differentiation was the primary organizing agent underlying the observed behavior. Secondly, the research we have cited involving the "release" of responses which have at least some of the phenotypic properties of the FAP, has employed only adult animals and consequently cannot rule out the possibility that the underlying neuronal configurations, instead of being genetically blueprinted, required experimentally instigated refinement before achieving the degree of functional specificity necessary to support response expression.

Finally, as both Gloor (1954) and Lehrman (1953, 1956) have emphasized, the instinct theorist has tended to overinterpret the work of Hess and his associates. The diencephalic areas they stimulated did not exhibit the characteristics of self-contained encapsulated units awaiting only adequate releasing stimuli before discharging their responses. Rather, the effects of stimulation depended partly on the nature of concurrent afferent inflow, the form of the response as well as its latency of release being modified accordingly. This, of course, does not corroborate the view that

autogenous differentiation is sufficient to provide for the coordination of the FAP. Indeed, whatever the extent of their genetically induced organization, it is possible that these areas must be continuously directed by information coming from cortical and somatic sources in order to support temporally integrated behavior. Whether the same interpretation can also be offered in relation to the work of von Holst and St. Paul remains to be determined.

Response Elicitability and Action-Specific Energy

That a temporally integrated behavioral pattern like the FAP should seem to exhibit an activity-specific rhythm of exhaustion and recovery has posed challenging problems for all students of behavior. As already indicated, this rhythmicity has led the instinct theorist to propose a reservoir model of motivation and, in addition, to conjecture that the underlying process must be central in origin and neurologically akin to the evidently endogenous activity shown by the cardiac ganglion of the lobster and the brain stem of the goldfish.

One can never be certain, however, even when working with isolated nerve-cell preparations, that the environment is not partly responsible for the pattern of excitation that is present (Adrian, 1950). All that one can be sure of is that a particular neural rhythm is not the result of a corresponding afferent rhythm. However, there is a great deal of evidence which, although not absolutely conclusive, nonetheless makes it reasonable to infer that the invertebrate as well as the vertebrate nervous system is continuously active, generating patterned impulse-sequences that arise neither from phasic nor from nonphasic sensory input (Bullock, 1961; Precht,

1956). But to grant the validity of this inference is not tantamount to consenting to the instinct theorist's attempt to relate such endogenously determined impulse sequences to fluctuations in the elicibility of the FAP. There are simply no neurological data which convincingly support the idea that the motivational dynamics of the FAP, or of any other temporally integrated movement pattern for that matter, are controlled and regulated by the automatic production of response-specific impulses.

To illustrate the inductive gap which in fact exists here, consider the type of evidence to which the instinct theorist has repeatedly appealed. Weiss (1941) implanted a piece of spinal cord and a limb bud onto the back of an amblystoma embryo and found that movement occurred as soon as efferent fibers made contact with the limb despite the fact that afferent connections had not yet been established. Although this result does provide dramatic evidence of neural automaticity, it certainly does not show that the source of energy for a response pattern as complex as the FAP is neurologically akin to the innervation of a limb bud nor even that coordinated locomotion can arise independently of sensory input. In relation to the question of coordination, it should be emphasized that the movements Weiss observed were only spasmodic or paroxysmal, hence, differing markedly from the smooth ambulatory rhythm normally evinced by the adult amblystoma. Significantly enough, studies by Gray and Lissman (1940, 1946a, 1946b) have demonstrated that peripheral excitation is absolutely essential for the typical locomotory pattern of the amphibian and, by implication, for all other vertebrates as well (also see Gray, 1950).

Of course, the status of the reser-

voir model of motivation does not rest solely on the neurological speculations advanced by the instinct theorist. The behavioral implications of the model are considerably more important and it is upon them that a definitive evaluation must be based.

The reservoir model makes reduction in elicibility contingent upon response expression since it holds that the discharge of action-specific energy can occur only through the performance or execution of the response. On this basis, it seems appropriate to ask whether reduction in elicibility necessarily requires response performance and, correlatively, whether a behavioral model involving the accumulation and release of action specific energy is generally applicable.

Prechtl (1953) found that the gaping response of young chaffinches can be elicited by visual, acoustic, and vibratory stimuli. He then discovered that the exhaustion of gaping by repeated presentation of one stimulus left the nestling relatively unaffected in its readiness to gape to the other stimuli. Hinde (1954, 1960a, 1960b) has also reported similar results in relation to other responses having the functional characteristics of the FAP, and has emphasized that, for such responses at least, stimulus-specific satiation rather than the discharge of response-specific energy appears to be the mechanism effecting reduction in elicibility.

The reservoir model must also be considered inadequate because it fails to incorporate the fact that FAP availability can change as a function of other types of afferent relationships which are perhaps even more subtle. A relevant case here is the incubation behavior of the black-headed gull (Beer, 1961, 1962; Moynihan, 1953). In settling on a clutch of eggs, these birds perform a sequence of move-

ments which include lowering the chest, waggling the tail from side to side, and rapidly quivering the body. Of particular significance as Beer has pointed out, is the fact that the frequency with which this sequence occurs is related not to energy consumption, but in part at least, to thermal effects consequent upon settling. More specifically, if the pattern of sensory stimulation received from the eggs fails to satisfy the temperature requirements of the brood patch, and probably the requirements of other structures as well, then the bird will begin again and repeat the entire settling sequence. In fact, it can be made to do this many times simply by decreasing the normal clutch number or by causing the eggs to become disarranged.

In addition to stimulus effects of the kind we have just discussed, brief mention might also be made of the role played by the absolute level of afferent input. Hebb (1955) has drawn attention to the significance of the brain-stem reticular system, proposing that its innervation by peripheral inflow induces a general arousal state which influences all behavior regardless of the specific control mechanisms involved. The reticular formation, functioning in the capacity of a non-specific activator, could then produce widespread threshold changes. If this in fact proves to be so in relation to the FAP, then the reservoir model would have to accommodate not only the influence of specific stimulus factors but the influence of general ones as well.

It seems clear from the evidence cited thus far that reduction in the elicibility of an FAP is not necessarily contingent on its performance. On this basis, it is reasonable to conclude that the reservoir model, involving as it does the concept of response-

effected energy discharge, does not offer a framework comprehensive enough to integrate all instances of threshold change in the FAP.

AN EPIGENETIC APPROACH TO THE FAP

The above analysis indicates that the theoretical position of the instinct theorist is not in fact demanded by the empirical properties of the FAP. Indeed, we are able to show that each of these properties can reasonably be interpreted in a manner broadly consonant with what may be called an epigenetic approach. In conclusion it would be well to attempt a general characterization of the major features of this approach.

An epigenetic approach holds that all response systems are synthesized during ontogeny and that this synthesis involves the integrative influence of both intraorganic processes and extrinsic stimulative conditions. It considers gene effects to be contingent on environmental conditions and regards the genotype as capable of entering into different classes of relationships depending on the prevailing environmental context. In the epigeneticist's view, the environment is not benignly supportive, but actively implicated in determining the very structure and organization of each response system.

Contrast this view with the template conception of the instinct theorist which maintains that FAPs result from the virtually passive translation of genetic factors into phenotypic traits through the medium of tissue growth and tissue differentiation. Each FAP, in other words, is represented in the genome by functionally congruent genic agents which become transmuted into underlying neural coordinating centers as long as conditions consonant with viability prevail.

According to an epigenetic concep-

tion of development, however, there is no adaptively organized response—species-typic or otherwise—that exhibits the genetic parallelism of which the instinct theorist speaks. Rather, the behavioral significance of genetic entities, and consequently of organic traits, is conceived to derive, in part at least, from the particular environmental context with which they become intermeshed during ontogeny. It is this coalescence of the endogenous and the extrinsic which, in an epigenetic view, makes each integrated response an emergent and which thus renders it gravely misleading to conceptualize maturational elements as functioning isomorphically in behavioral development.

REFERENCES

- ADRIAN, E. D. The control of nerve-cell activity. *Symposia of the Society for Experimental Biology*, 1950, 4, 85-91.
- ADRIAN, E. D., & BUYTENDIJK, F. J. J. Potential changes in the isolated brain of the goldfish. *Journal of Physiology*, 1931, 71, 121-130.
- ADRIAN, E. D. Potential changes in the isolated nervous system of *Dytiscus marginalis*. *Journal of Physiology*, 1931, 72, 132-151.
- ANDERSSON, B. The effect of injections of hypertonic NaCl-solutions into different parts of the hypothalamus in goats. *Acta Physiologica Scandinavica*, 1953, 28, 188-201.
- BAERENDS, G. P. The ethological analysis of fish behavior. In M. E. Brown (Ed.), *The physiology of fishes*. Vol. 2. New York: Academic Press, 1957. Pp. 229-269.
- BAERENDS, G. P., & BAERENDS-VON ROON, J. M. An introduction to the study of the ethology of cichlid fishes. *Behaviour*, Supplement I, 1950.
- BEACH, F. A. The descent of instinct. *Psychological Review*, 1955, 62, 401-410.
- BEER, C. G. Incubation and nest-building behaviour of black-headed gulls. I: Incubation behaviour in the incubation period. *Behaviour*, 1961, 18, 62-106.
- BEER, C. G. Incubation and nest-building behaviour of black-headed gulls. II: In-

- cubation behaviour in the laying period. *Behaviour*, 1962, 19, 283-304.
- BIRCH, H. G. Sources of order in the maternal behavior of animals. *American Journal of Orthopsychiatry*, 1956, 26, 279-284.
- BLEST, A. D. The evolution, ontogeny and quantitative control of the settling movements of some new world saturniid moths, with some comments on distance communication by honey-bees. *Behaviour*, 1960, 16, 188-253.
- BROWN, F. A., JR. The rhythmic nature of animals and plants. *American Scientist*, 1959, 47, 147-168.
- BULLOCK, T. H. The origins of patterned nervous discharge. *Behaviour*, 1961, 17, 48-59.
- CHOW, K. L., & NISSEN, H. W. Interocular transfer of learning in visually naive and experienced infant chimpanzees. *Journal of Comparative and Physiological Psychology*, 1955, 48, 229-237.
- CRANE, J. Eastern Pacific Expeditions of the New York Zoological Society. XXVI. Crabs of the genus *Uca* from the West Coast of Central America. *Zoologica*, 1941, 26, 145-208.
- CRANE, J. Basic patterns of display in fiddler crabs (*Ocypodidae*, Genus *Uca*). *Zoologica*, 1957, 42, 69-82.
- CULLEN, E. Experiment on the effect of social isolation on reproductive behaviour in the three-spined stickleback. *Animal Behaviour*, 1960, 8, 235. (Abstract)
- DAVIS, J. Comparative foraging behaviour of the spotted and brown towhees. *Auk*, 1957, 74, 129-166.
- DILGER, W. Nest material carrying behavior of F_1 hybrids between *Agapornis fischeri* and *A. roseicollis*. *Anatomical Record*, 1959, 134, 554.
- DILGER, W. C. Behavior and genetics. In E. C. Bliss (Ed.), *Roots of behavior*. New York: Harper, 1962. Pp. 35-47.
- EIBL-EIBESFELDT, I. The interactions of unlearned behaviour patterns and learning in mammals. In J. F. Delafresnaye (Ed.), *Brain mechanisms and learning*. Oxford: Blackwell Scientific Publications, 1961. Pp. 52-73.
- EIBL-EIBESFELDT, I., & KRAMER, S. Ethology, the comparative study of animal behavior. *Quarterly Review of Biology*, 1958, 33, 181-211.
- EVANS, H. E. Comparative ethology and the systematics of spider wasps. *Systematic Zoology*, 1953, 3, 155-172.
- FULLER, J. L., & THOMPSON, W. R. *Behavior genetics*. New York: Wiley, 1960.
- GLOOR, P. Autonomic functions of the diencephalon. A summary of the experimental work of Prof. W. R. Hess. *Archives of Neurological Psychiatry*, 1954, 71, 773-790.
- GRAY, J. The role of peripheral sense organs during locomotion in vertebrates. *Symposia of the Society of Experimental Biology*, 1950, 4, 112-126.
- GRAY, J., & LISSMAN, H. W. The effect of de-afferentation upon the locomotor activity of amphibian limbs. *Journal of Experimental Biology*, 1940, 17, 227-236.
- GRAY, J., & LISSMAN, H. W. Further observations on the effect of de-afferentation on the locomotor activity of amphibian limbs. *Journal of Experimental Biology*, 1946, 23, 121-132. (a)
- GRAY, J., & LISSMAN, H. W. Coordination of limb movements in the amphibia. *Journal of Experimental Biology*, 1946, 23, 133-142. (b)
- HARKER, J. E. Diurnal rhythms in the animal kingdom. *Biological Reviews and Biological Proceedings of the Cambridge Philosophical Society*, 1958, 33, 1-52.
- HARRIS, G. W., MICHAEL, R. P., & SCOTT, P. P. Neurological site of action of stilbesterol in eliciting sexual behavior. In *Neurological basis of behavior*. Boston: Little, Brown, 1958. Pp. 236-254.
- HEBB, D. O. On the nature of fear. *Psychological Review*, 1946, 53, 250-275.
- HEBB, D. O. Heredity and environment in mammalian behaviour. *British Journal of Animal Behaviour*, 1953, 1, 43-47.
- HEBB, D. O. Drives and the C.N.S. (Conceptual nervous system). *Psychological Review*, 1955, 62, 243-254.
- HESS, W. R. Das Zwischenhirn: Syndrome, Lokalisationem, Funktionem. Basel: Benno Schwabe, 1949.
- HESS, W. R., & BRUGGER, M. Das subkortikale Zentrum der Affektiven Abwehrreaktion. *Helvetica physiologica et pharmacologica acta*, 1943, 1, 33-52.
- HINDE, R. A. Changes in responsiveness to a constant stimulus. *British Journal of Animal Behaviour*, 1954, 2, 41-55.
- HINDE, R. A. A comparative study of the behaviour of certain finches. *Ibis*, 1955, 97, 706-745, 98, 1-23.
- HINDE, R. A. The behaviour of certain cardueline F_1 interspecies hybrids. *Behaviour*, 1956, 9, 202-213.
- HINDE, R. A. Some recent trends in ethology. In S. Koch (Ed.), *Psychology: A study of a science*. Vol. 2. New York: McGraw-Hill, 1959. Pp. 561-610.

- HINDE, R. A. Concepts of drive. In M. Brazier (Ed.), *The central nervous system and behavior*. New York: Josiah Macy Foundation, 1960. Pp. 277-305. (a)
- HINDE, R. A. Motivation. *Ibis*, 1960, 101, 353-357. (b)
- HINDE, R. A., & TINBERGEN, N. The comparative study of species-specific behavior. In A. Roe & G. Simpson (Eds.), *Behavior and evolution*. New Haven: Yale Univ. Press, 1958. Pp. 251-268.
- HOLST, E., VON. Studien über Reflexe und Rhythmen beim Goldfisch (*Carassius auratus*). *Zeitschrift für vergleichende Physiologie*, 1934, 20, 582-599.
- HOLST, E. VON, & ST. PAUL, U. Von Wirkungsgefüge der Triebe. *Naturwissenschaften*, 1960, 18, 409-422.
- HOLST, E. VON, & ST. PAUL, U. Electrically controlled behavior. *Scientific American*, 1962, 206, 50-59.
- IERSEL, J. J. A., VAN. An analysis of the parental behaviour of the male three-spined stickleback. *Behaviour*, Suppl. III, 1953.
- LEHRMAN, D. S. A critique of Konrad Lorenz's theory of instinctive behavior. *Quarterly Review of Biology*, 1953, 28, 337-363.
- LEHRMAN, D. S. On the organization of maternal behavior and the problem of instinct. In *L'instinct dans le comportement des animaux et de l'homme*. Paris: Masson et Cie, 1956. Pp. 475-514.
- LORENZ, K. Über die Bildung Instinkt-begriffes. *Die Naturwissenschaften*, 1937, 25, 289-300, 307-318, 324-331. Translated in C. Schiller (Ed.), *Instinctive behavior*. New York: International Universities Press, 1957. Pp. 129-175.
- LORENZ, K. Morphology and behavior patterns in closely allied species. In B. Schaffner (Ed.), *Group processes*. New York: Josiah Macy Foundation, 1955. Pp. 75-163.
- LORENZ, K. The objectivistic theory of instinct. In *L'instinct dans le comportement des animaux et de l'homme*. Paris: Masson et Cie, 1956. Pp. 51-64.
- LORENZ, K. The evolution of behavior. *Scientific American*, 1958, 199, 67-73.
- LORENZ, K. Prinzipien der Vergleichenden Verhaltensforschung. *Fortschritte der Zoologie*, 1960, 12, 265-294. Translated by E. Klinghammer, University of Chicago.
- LORENZ, K. Phylogenetische Anpassung und Adaptive Modifikation des Verhaltens. *Zeitschrift für Tierpsychologie*, 1961, 18, 139-187.
- LORENZ, K., & TINBERGEN, N. Taxis und Instinkthandlung in der Eirollbewegung der Graugans. *Zeitschrift für Tierpsychologie*, 1938, 2, 1-29. Translated in C. Schiller (Ed.), *Instinctive behavior*. New York: International Universities Press, 1961. Pp. 176-208.
- MAIER, N. R. F., & SCHNEIRLA, T. C. *Principles of animal psychology*. New York: McGraw-Hill, 1935.
- MAYNARD, D. M. Activity in a crustacean ganglion. II: Pattern and interaction in burst formation. *Biological Bulletin*, 1955, 109, 420-436.
- MAYR, E. Behavior and systematics. In A. Roe & G. Simpson (Eds.), *Behavior and evolution*. New Haven: Yale Univ. Press, 1958. Pp. 341-362.
- MCBRIDE, A. F., & HEBB, D. O. Behavior of the captive bottle-nose dolphin *Tursiops truncatus*. *Journal of Comparative and Physiological Psychology*, 1948, 41, 111-123.
- MELZACK, R. Irrational fears in the dog. *Canadian Journal of Psychology*, 1952, 6, 141-147.
- MELZACK, R., & SCOTT, T. H. The effects of early experience on the response to pain. *Journal of Comparative and Physiological Psychology*, 1957, 50, 155-161.
- MILLER, N. E. Experiments on motivation: studies concerning psychological, physiological and pharmacological techniques. *Science*, 1957, 126, 1271-1278.
- MILLER, N. E. Central stimulation and other new approaches to motivation and reward. *American Psychologist*, 1958, 13, 100-108.
- MILLER, N. E. Motivational effects of brain stimulation and drugs. *Federation Proceedings*, 1960, 19, 846-854.
- MOLTZ, H. The fixed action pattern: empirical properties and theoretical implications. In J. Wortis (Ed.), *Recent advances in biological psychiatry*. Vol. 4. New York: Plenum Press, 1962. Pp. 69-85.
- MOYNIHAN, M. Some displacement activities of the Black-headed Gull. *Behavior*, 1953, 5, 58-80.
- PRECHTL, H. F. R. Zur Physiologie der Angeborenen Auslösenden Mechanismen. I. Quantitative Untersuchungen über die sperrbewegung junger Singvogel. *Behaviour*, 1953, 5, 32-50.
- PRECHTL, H. F. R. Neurophysiologische Mechanismen des formstarren Verhaltens. *Behaviour*, 1956, 9, 243-319.
- PRECHTL, H. F. R. The directed head turning response and allied movements of the

- human baby. *Behaviour*, 1958, 13, 212-242.
- RAMSAY, A. O. Behaviour of some hybrids in the mallard group. *Animal Behaviour*, 1961, 9, 105-106.
- RIESEN, A. H. Post-partum development of behavior. *Chicago Medical School Quarterly*, 1951, 13, 17-24.
- RIESEN, A. H. Plasticity of behavior: Psychological series. In H. F. Harlow & C. N. Woolsey (Eds.), *Biological and biochemical bases of behavior*. Madison: Univer. Wisconsin Press, 1958. Pp. 425-450.
- RIESEN, A. H. Receptor functions. In P. H. Mussen (Ed.), *Handbook of research: Methods in childhood development*. New York: Wiley, 1960.
- RIESEN, A. H. Stimulation as a requirement for growth and function in behavioral development. In D. W. Fiske & S. R. Maddi (Eds.), *Functions of varied experience*. Homewood, Ill.: Dorsey Press, 1961. Pp. 57-80.
- SCHNEIRLA, T. C. Problems in the biopsychology of social organization. *Journal of Abnormal and Social Psychology*, 1946, 41, 385-402.
- SCHNEIRLA, T. C. Interrelationships of the "innate" and the "acquired" in instinctive behavior. In *L'instinct dans le comportement des animaux et de l'homme*. Paris: Masson et Cie, 1956. Pp. 387-439.
- SCHNEIRLA, T. C. The concept of development in comparative psychology. In D. B. Harris (Ed.), *The concept of development: An issue in the study of human behavior*. Minneapolis: Univer. Minnesota Press, 1957. Pp. 78-108.
- SCHNEIRLA, T. C., & ROSENBLATT, J. Behavioral organization and genesis of the social bond in insects and mammals. *American Journal of Orthopsychiatry*, 1961, 31, 223-253.
- SPERRY, R. W. Mechanisms of neural maturation. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951.
- SPERRY, R. W. Physiological plasticity and brain circuit theory. In H. F. Harlow & C. N. Woolsey (Eds.), *Biological and biochemical bases of behavior*. Madison: Univer. Wisconsin Press, 1958.
- STEINBERG, JUNE, & BINDRA, D. Effects of pregnancy and salt-intake on genital licking. *Journal of Comparative and Physiological Psychology*, 1962, 55, 103-106.
- THORPE, W. H. Comparative psychology. *Annual Review of Psychology*, 1961, 12, 27-50.
- THORPE, W. H. *Learning and instinct in animals*. Cambridge: Harvard Univer. Press, 1956.
- THORPE, W. H. Some concepts of ethology. *Nature, London*, 1954, 174, 101-106.
- THORPE, W. H., & JONES, F. G. W. Olfactory conditioning in a parasitic insect and its relation to the problem of host selection. *Proceedings of the Royal Society of London, Series B*, 1937, 124, 56-81.
- TINBERGEN, N. An objectivistic study of the innate behaviour of animals. *Bibliotheca biotheoretica, Leiden*, 1942, 1, 39-98.
- TINBERGEN, N. *The study of instinct*. Oxford: Oxford Univer. Press, 1951.
- TINBERGEN, N. The origin and evolution of courtship and threat display. In A. C. Hardy, J. S. Huxley, & E. B. Ford (Eds.), *Evolution as a process*. London: Allen & Unwin, 1954.
- TINBERGEN, N. Some aspects of ethology, the biological study of animal behaviour. *British Association for the Advancement of Science*, 1955, 12, 17-19.
- TINBERGEN, N. Comparative studies of the behaviour of gulls (Laridae): A progress report. *Behaviour*, 1959, 15, 1-70.
- TINBERGEN, N., & KUENEN, D. J. Über die auslösenden und die richtunggebenden Reizsituationen der Sperrbewegung von jungen Drosseln. *Zeitschrift für Tierpsychologie*, 1939, 3, 37-60. Translated in C. Schiller (Ed.), *Instinctive behavior*. New York: International Universities Press, 1961. Pp. 209-238.
- TINBERGEN, N., & PERDECK, A. C. On the stimulus situation releasing the begging response of the newly hatched herring gull chick (*Larus argentatus argentatus*). *Behaviour*, 1951, 3, 1-39.
- VAN DER KLOOT, G., & WILLIAMS, C. M. Cocoon construction of the cecropia silkworm. I: The role of the external environment. *Behaviour*, 1953, 5, 141-156. (a)
- VAN DER KLOOT, G., & WILLIAMS, C. H. Cocoon construction of the cecropia silkworm. II: The role of the internal environment. *Behaviour*, 1953, 5, 157-174. (b)
- VERPLANCK, W. S. Since learned behavior is innate, and vice versa, what now? *Psychological Review*, 1955, 62, 139-144.
- WEISS, P. Self-differentiation of the basic patterns of coordination. *Comparative Psychology Monographs*, 1941, 17, 1-96.

(Received November 1, 1963)

SCOPE AND GENERALITY OF VERBALLY DEFINED PERSONALITY FACTORS¹

DONALD R. PETERSON

University of Illinois

Factor analyses of verbal personality measures have typically generated highly complex multidimensional structural systems. Available evidence now suggests that the most dependable dimensions drawn from conventional factor analyses of ratings and questionnaires are simple, familiar dimensions of broad semantic scope. It also appears that most of the initially obscure, apparently more precise, more narrowly defined factors many investigators claim to have revealed are either trivial, artifactual, capricious, or all 3. Verbal descriptions of personality were reduced to 2 factors, and the 2 factors were reduced to 2 ratings, 1 concerning perceived adjustment and the other related to introversion-extraversion. Convergent and discriminant validities for the 2 simple ratings are evidently equal or superior to those for any of the more cumbersome and expensive questionnaires and rating schedules examined in the review. Implications for theory and method were discussed.

Of the several decisions required in the conduct of factor analysis, none has greater bearing on results than the number of factors to keep for rotation. When only a few "strong" factors are retained, the dimensions which emerge are broad in scope. Each is defined by a large set of variables, which may appear semantically heterogeneous. When many factors are retained, the dimensions which emerge are narrow in scope. Each is defined by a relatively small number of variables. These may also be semantically heterogeneous. When a few factors of high variance are rotated against a larger number of factors of low variance, the former are rather narrowly fractionated and often lose their original identity. But if the many narrow factors are obliquely determined, and a second-order analysis is conducted, the original massive factors are essen-

tially reconstituted in the resulting second-order dimensions.

Insofar as the smaller factors are more than spurious, we are dealing with an issue of factor "order" or degree of complexity, and mathematically there is neither need nor method to make a choice. The factors simply vary in complexity. In general, any narrow factors can be merged into broader factors and any broad factors can be separated into narrower factors. With appropriate manipulation of measures, this can be continued indefinitely in either direction, though within the limits of a given matrix, the process has to stop when a single general factor emerges at the top of the hierarchy, and instrumental specifics appear at the bottom (Burt, 1950; Humphreys, 1962; Vernon, 1950).

In the pursuit of research and in the construction of theory, however, choice of level is often desirable. The usual criteria for factor inclusion lead to the retention of large numbers of

¹ An earlier version of this paper was read at meetings of the Society for Research in Child Development, Berkeley, California, April 1963.

factors and a complex structural system. It is the purpose of this article to suggest that better description can usually be attained with a very few factors, measured very simply. For the time being, conclusions are restricted to conventional factor analyses of verbal personality descriptions, where the goals are to define dimensions of personality, the data are drawn from ratings and questionnaires, and the analyses are conducted within the bounds of a communality, simple structure model. The comments to follow may raise questions about other domains, other media, and other models, but without further data they can do no more than that.

DESCRIPTIVE EFFICIENCY

Factor analysis was designed to achieve parsimony in the description of observed phenomena. Descriptive efficiency is therefore of primary concern in evaluating the utility of a structural system.

In texts and theoretical articles, variance plots for succeeding factors are ordinarily represented by a smoothly accelerated, gradually descending function of the sort shown by the broken line in Figure 1. If this were actually the form of the curve, nonarbitrary decisions about factor inclusion would have to be based on some other consideration than the amounts of variance accommodated by successive factors. But this is not the way the function usually looks. In a number of empirical investigations, it has taken the form shown by the solid line in Figure 1. While Cattell and other theorists might expect the second factor to take up almost as much variance as the first, the second factor actually accounts for less than half as much variance as the first. There is a sharp rise in acceleration after the first two or three factors and very low slope

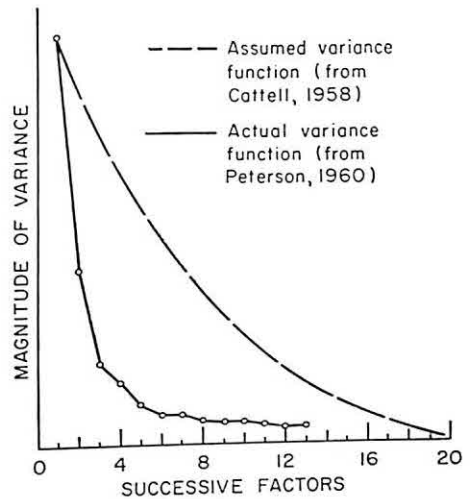


FIG. 1. Assumed and actual variance functions.

after that. Under such conditions, it is inefficient to retain any factors beyond the first few.

If one goes beyond that, it is difficult to decide where to stop. The addition of new factors and the division of old ones seems to be endless, and I believe we are already near the point of utter descriptive chaos in the proliferation of personality dimensions. In the field of human abilities, Guilford now has more factors than Thurstone had tests (Humphreys, 1962). The situation in the personality domain is even more terrifying. Cattell (1957a) alone has defined at least 45 dimensions. Those are the ones he mentioned in his 1957 book and he left some out. Extracting large numbers of factors in the face of facts as shown in Figure 1 defeats the purpose of factor analysis. If an investigator wants to deal with factors at all, he had better deal with factors which are reasonably efficient. If there is need to deal with more specific characteristics, factorization offers no advantage. The number of dimensions one must eventually consider will be enormously large. Researchers might

TABLE 1
DISTRIBUTIONS OF PRINCIPAL AXIS FACTOR LOADINGS

Range	Empirical ratings Factor number										
	1	2	3	4	5	6	7	8	9	10	11
.80 to 1.00	1	—	—	—	—	—	—	—	—	—	—
.60 to .79	5	5	1	—	—	—	—	—	—	—	—
.40 to .59	1	3	—	3	—	—	—	—	1	—	—
.20 to .39	3	5	4	3	3	2	5	5	1	1	2
.00 to .19	2	5	15	13	15	12	10	19	12	12	13
-.20 to -.01	4	8	11	15	16	18	17	10	21	18	18
-.40 to -.21	6	5	4	2	1	3	4	2	—	5	3
-.60 to -.41	4	3	1	—	1	1	—	—	1	—	—
-.80 to -.61	7	2	—	—	—	—	—	—	—	—	—
-1.00 to -.81	3	—	—	—	—	—	—	—	—	—	—

Range	Random variables Factor number										
	1	2	3	4	5	6	7	8	9	10	11
.80 to 1.00	—	—	—	—	—	—	—	—	—	—	—
.60 to .79	—	—	—	—	—	—	—	—	—	—	—
.40 to .59	—	—	2	1	2	2	—	—	—	—	—
.20 to .39	2	1	5	1	2	3	5	1	2	2	2
.00 to .19	8	5	7	9	3	7	8	7	13	4	9
-.20 to -.01	9	12	3	6	9	7	4	9	3	11	6
-.40 to -.21	1	2	3	2	4	1	3	2	2	3	3
-.60 to -.41	—	—	—	1	—	—	—	1	—	—	—
-.80 to -.61	—	—	—	—	—	—	—	—	—	—	—
-1.00 to -.81	—	—	—	—	—	—	—	—	—	—	—

as well examine the original elemental variables, and not bother with factor analysis at all.

SUFFICIENCY IN FACTOR EXTRACTION

Not only is the retention of many weak factors inefficient, at least for some kinds of data the smaller dimensions appear to be purely artifactual. Linn (1964) introduced 20 random variables among 36 other "true" variables which had previously been examined in a study of rated personality characteristics (Peterson & Cattell, 1959). Results are shown in Table 1. The distributions of loadings for random variables become indistinguishable from those for "real" variables

when the number of factors exceeds two or three. Mean values derived from Table 1 show that the mean absolute loading for random variables actually exceeds that for empirical ratings for the third factor, and the mean loadings are approximately the same for random as for "real" variables for all factors beyond the third. Retention of 13 factors, as was done in the original research, is obviously out of the question.

FACTOR INVARIANCE

By "invariance" is meant a condition of factor comparability such that factors in two or more studies remain identifiably the same when the con-

ceptual domain is unchanged; the general rules for accepting representative measures within that domain are unchanged, but the actual measures employed are altered in some way, and/or the subject sample is changed. Invariance, in this sense, is important to attain in any factor analysis, and is particularly important in the use of factor analysis as a means of defining general traits.

Given replication, the invariance of factors can be examined in a number of ways. The most common procedure involves a combination of inspection, intuition, and preconception. The investigator gazes thoughtfully at the salient variables, and decides whether one factor "appears" to be the same as another and should be labeled accordingly. The criteria are seldom explicit, and weights and interactions are unknown.

Efforts to objectify and standardize procedures for assessing loading profile similarity have taken several forms (Cattell, 1955; Cattell & Baggailey, 1960; Harman, 1960, pp. 256-259; Tucker, 1951). The data shown below were derived by intercorrelating columns of factor loadings over common marker variables. In relation to the usual sort of correlation, the factors are analogous to variables, and loadings on common marker items are analogous to scores obtained by subjects. If all the columns of loadings in one matrix are correlated with all the columns of loadings in another, a rectangular matrix can be formed, with the factors in one study heading the rows, and factors in the other heading the columns. If the factors have been previously identified, by inspection, elimination, intuition, or otherwise, it is possible to place the factors in the two studies in parallel order. Then, if allegedly matching factors are the same or similar in fact, high r 's should

appear in the diagonal, and low values everywhere else.

Over the past several years, with the help of several colleagues and a very effective computer system, the author has evaluated the invariance of both broad and narrow factors from ratings of personality in children, ratings of clinical problems in children, ratings of parental attitudes, and questionnaire responses which discriminate between delinquent and nondelinquent adolescents. In every case, considering each set of factors as a whole, the broad group factors have shown greater stability from one study to another. The following example is chosen principally on the basis of adequacy of both studies, freedom from contamination of results, and only incidentally because the findings are slightly more favorable to the author's case than some of the other comparisons which might have been chosen.

Some time ago, Cattell (1947) obtained ratings of young adults, analyzed the data, and obtained 11 factors which he identified and named. Later, analogous ratings of 7-year-old children were obtained, a similar analysis was performed, and 13 factors were identified. For the most part, these "appeared" to resemble the factors for adults, and the authors concluded that most of the factors previously isolated in older subjects were found for the younger ones (Cattell & Coan, 1957).

Intercorrelation of loadings, however, produced the matrix which appears as Table 2. The correlations in the diagonal are somewhat higher than those in the off-diagonal cells. This is not surprising, since loading profile is the main basis for asserting identity in the first place. But the r 's for allegedly matching factors are not impressively higher than for nonmatching factors. Nearly every factor is as

TABLE 2
CORRELATIONS BETWEEN "NARROW" FACTOR LOADINGS FOR ADULTS
AND FOR 7-YEAR-OLD CHILDREN

	Factor	Adults										
		A	E	F	G	H	I	J	L	M	B	K
Seven-year-olds	A	.08	.06	.32	.03	.32	.20	.05	.18	.51	.16	.23
	E	.49	.32	.33	.21	.29	.10	.14	.07	.24	.31	.01
	F	.21	.03	.49	.02	.13	.12	.41	.02	.28	.08	.07
	G	.35	.32	.19	.61	.24	.04	.28	.24	.20	.50	.00
	H	.01	.07	.29	.25	.70	.08	.18	.14	.04	.27	.36
	I	.23	.07	.06	.32	.16	.41	.13	.45	.01	.03	.21
	J	.56	.07	.01	.22	.01	.06	.37	.17	.07	.10	.25
	L	.18	.11	.11	.18	.13	.19	.17	.40	.03	.22	.15
	M	.14	.01	.28	.21	.22	.07	.01	.23	.14	.31	.39
	C	.15	.15	.11	.13	.11	.02	.13	.17	.13	.02	.03
	D	.11	.11	.15	.11	.04	.35	.09	.08	.22	.05	.17
	L'	.09	.09	.13	.10	.17	.13	.01	.49	.16	.13	.16
	O	.26	.27	.19	.27	.00	.10	.23	.19	.33	.07	.10

well matched with another from which it is presumably distinct as with the one it is said to resemble.

When the two high order dimensions, appearing to reflect "adjustment" and "introversion-extraversion" were examined in the same way, the figures in Table 3 emerged. Correlations in the main diagonal, for matching factors, are high. The other correlations, for nonmatching factors, are low.

Data from other comparisons are summarized in Table 4. In this table, the values for matching factors are the means of diagonal values from matrices

TABLE 3
CORRELATIONS BETWEEN "BROAD" FACTOR
LOADINGS FOR ADULTS AND FOR
7-YEAR-OLD CHILDREN

		7-year-olds	
		Adjustment	Extraversion-introversion
Adults	Adjustment	.92	.27
	Extraversion-introversion	.10	.83

of the kind shown in Tables 2 and 3. The values for nonmatching factors are the means of all other correlations in each matrix, taken without regard to sign.

The correlations in the first column of Table 4, for similar "strong" factors, tend to fall in the 80s, and are higher in each case than the averages for factors alleged to be different. The figures in the first column are also higher than those in the third column, which contains similarity indexes for narrow factors which theoretically ought to match. In fact the latter values are not much higher than those in the fourth column, for the many factors which are hypothetically distinct from one another.

Factor invariance in studies of this sort may arise from covariance in observed behavioral tendencies, as is ordinarily assumed by trait theorists. Or it may have more to do with the perceptual tendencies of observers, as will be proposed later in this article. Demonstrations of invariance say nothing to this point. But structural replicability is a high desideratum in

TABLE 4

AVERAGE FACTOR LOADING CORRELATIONS FOR BROAD AND NARROW FACTORS

	Broad factors		Narrow factors	
	Mean r for matching factors	Mean r for nonmatching factors	Mean r for matching factors	Mean r for nonmatching factors
Personality factors				
Adults (Cattell, 1947)				
versus 11-year-olds	.89	.26	.27	.18
(Cattell & Gruen, 1953)				
Adults (Cattell, 1947)				
versus 7-year-olds	.82	.18	.40	.17
(Cattell & Coan, 1957)				
Adults (Cattell, 1947)				
versus 4-year-olds	.80	.32	.49 ^a	.18
(Peterson & Cattell, 1959)				
Parental attitudes				
American mothers				
versus Sicilian mothers	.72	.18	.40	.18
Children's problems (Peterson, 1961)				
Fifth and sixth grades				
versus third and fourth grades	.82	.41	.55 ^b	.22
Fifth and sixth grades				
versus first and second grades	.84	.49	.51 ^b	.30
Fifth and sixth grades				
versus kindergarten	.84	.41	.52 ^b	.26

^a This value involves maximal capitalization on chance. Matching was done after the loading correlations were known.

^b Narrow factors in these three groups were defined in matrices containing only 5 factors, instead of the 10-15 factors in the other comparisons of narrow factors. This probably explains the improvement in results.

itself, and the evidence clearly favors a few strong factors over many weak ones in this regard.

VALIDITY OF VERBAL PERSONALITY MEASURES

If questionnaires and ratings used in personality measurement are indeed measuring personality, the traits involved should manifest some constancy over methods, over situations, to some extent over time, and in the case of ratings should display appreciable stability with shifts in the identity of raters. If conditions of mini-

mal validity are to be met, the various intercorrelations between measures of a given trait ought to be higher than those between different traits as assessed by separate measures (Campbell & Fiske, 1959). It is also desirable, if not absolutely essential, for the correlations between measures of a given trait to exceed those within methods and situations (Humphreys, 1960).

It is here, as Becker (1960) has pointed out, that the multifactor measures of Cattell have failed most badly. In one of the studies Becker cites as

TABLE 5

RELATIONSHIPS BETWEEN Q FACTORS AND PARENT RATINGS OF PARALLEL FACTORS
Q FACTORS (Self-Ratings)

Parent ratings	A	B	C	E	F	G	H	I	L	M	N	O	Q ₁	Q ₂	Q ₃	Q ₄
A	.19	.04	-.14	.06	.02	.12	-.22	-.07	-.22	-.05	-.08	.04	-.05	.06	.08	.14
B	-.03	.03	.06	.08	-.05	.00	-.07	.31	.23	.02	-.15	.20	.06	-.12	-.03	.20
C	.15	.22	.27	-.23	.09	.13	.01	-.02	-.22	-.22	-.07	-.25	-.03	-.08	.16	-.27
E	.14	.08	-.06	.37	.10	-.21	.42	-.04	.13	.21	-.01	.06	.18	.30	-.08	.16
F	.32	.09	-.07	.22	.26	.05	.36	.12	-.25	.00	-.21	-.14	.04	.07	-.08	-.03
G	.14	.17	.15	-.29	.07	.18	-.02	-.05	-.17	-.29	.08	-.15	-.10	-.23	.14	-.20
H	.26	.26	.09	.14	.18	-.15	.45	-.12	.01	-.03	-.03	-.17	.11	.05	-.04	-.08
I	.06	-.06	-.15	.27	.01	-.07	.22	-.17	.22	.23	.07	.13	.03	.09	-.06	.23
L	-.28	.09	.06	.15	-.24	.01	.04	-.06	.14	.04	.09	-.09	.13	.21	.11	.13
M	-.07	-.05	-.09	.08	-.31	.01	-.08	-.14	.13	.05	-.01	.14	.02	.04	.02	.22
N	.09	.06	.28	-.02	.01	-.07	.10	-.23	-.24	-.15	.07	-.18	-.17	-.06	-.01	-.24
O	-.04	-.18	-.19	.02	-.02	.01	.07	-.06	.02	.19	-.03	.02	.08	.01	-.09	.22
Q ₁	-.12	.22	.02	.34	.03	-.12	.05	.22	.11	.15	-.02	.07	.05	.20	-.10	.09
Q ₂	-.10	.00	.04	.02	-.06	-.19	-.11	-.11	.15	.11	.08	.03	.20	.01	-.09	.07
Q ₃	.30	.11	.25	-.19	.14	.14	-.12	-.08	-.16	-.22	-.06	-.16	-.15	-.24	.16	-.40
Q ₄	-.14	-.17	-.17	.25	-.22	.08	-.02	-.12	.27	.02	.11	.18	.24	.27	.06	.35

most favorable to Cattell's position (Meeland, 1952), the average correlation for "matched" behavior rating and questionnaire factors is .37, and the average for "nonmatching" factors is .32. Becker found other studies even less convincing, and concluded that claims of correspondence between questionnaire and rating factors were poorly supported by available evidence. Nothing in Cattell's (1961) reply changed any of the facts.

Whether higher order factors will offer any improvement in validity is not well determined at this time. The writer knows of only one study in which factors at both levels were examined for factor correspondence over

media. Wetzel (1963) administered the 16 P-F Test (Cattell, 1957b) to a class of undergraduate psychology students, and obtained peer and parent ratings on the dimensions the questionnaire is supposed to measure. Interrelationships were established for 72 subjects. Parents ratings were made by the mothers for all cases where mother ratings were available ($N = 61$). The remaining 11 parent ratings were made by fathers. Peer ratings were obtained from other students, elected by the subject's themselves as "people who knew them well in the university setting, and could give reasonably accurate evaluations." Most of the subjects asked their roommates to rate them.

Illustrative findings for narrow factors are shown in Table 5. The validity correlations in the diagonal are not much different from the remaining correlations in the matrix.

Values for two second order factors, Dynamic Integration (self-perception of adjustment) and Introversion-Extraversion, are shown in Table 6. In this case, the correlations in the diagonal for the "same" factors are at least statistically significant, and the off-diagonal values are not.

TABLE 6
RELATIONSHIPS BETWEEN QUESTIONNAIRE
FACTORS AND RATINGS OF
PARALLEL FACTORS

Parent ratings		Questionnaire factors (self-ratings)	
		Dynamic integration	Introversion-extraversion
	Dynamic integration	.28	.05
	Introversion-extraversion	.10	.38

TABLE 7
AVERAGE VALIDITY CORRELATIONS FOR BROAD AND NARROW FACTORS

	Broad factors		Narrow factors	
	Mean r for "same" factors	Mean r for "different" factors	Mean r for "same" factors	Mean r for "different" factors
Q factors versus peer ratings	.31	.19	.22	.13
Q factors versus parent ratings	.33	.08	.17	.12
Peer ratings versus parent ratings	.25	.04	.17	.12

Results from other comparisons in the study are summarized in Table 7. Overall comparisons among the original multitudinous primary factors, the mean r for "matching" factors is .19, which is not reliably different from zero for an N of 72. The mean r for "nonmatching" factors is .12, which is smaller than the mean correlation for theoretically related dimensions, but not much. The mean r for matching second order factors is .30, which exceeds the usual limit of statistical significance, while that for hypothetically unrelated factors is .10. The essential criteria of dimensional identity proposed by Campbell and Fiske are therefore poorly met by the original narrow factors. The broad second order factors look somewhat better.

ECONOMY AND SIMPLICITY IN THE ASSESSMENT OF VERBAL BEHAVIOR

Let me now press the reductionism of the present argument one step further. With considerable dependability, "primary" factors of the sort Cattell and others postulate merge into second-order factors, which possess certain advantages over the primary factors. The broad group factors appear to represent rather diffuse concepts of

self and others. Perhaps it is possible to specify these concepts within closer topical limits, and reduce some of the most important aspects of verbal personality assessment to a couple of questions. As to theory, it is proposed that when people consider their own "personalities" or those of others, their statements revolve largely about such concepts as adjustment and extraversion-introversion. As to procedure, it is proposed that a good deal of the information people convey in verbal statements about themselves or others can be gained by asking them (a) "How well adjusted is X?" (where X may be the self or another person), and (b) "Where does X stand in regard to introversion-extraversion?"

Wetzel (1963) also asked all respondents, the core subjects themselves as well as peers and parents, to provide ratings on "adjustment" and "extraversion-introversion," each on a simple 7-point rating scale. These were correlated with the second-order factor scores which had been painstakingly calculated from performance on the 16 P-F test and from the rating schedules. The correlation between the rating of adjustment and the second-order factor "Dynamic Integration versus

TABLE 8

MULTITRAIT-MULTIMETHOD CORRELATIONS FOR TWO SIMPLE PERSONALITY RATINGS

Self				Self				Peer			
		Adjust- ment	Intro- version- extra- version			Adjust- ment	Intro- version- extra- version			Adjust- ment	Intro- version- extra- version
Peer	Adjust- ment	.24	-.09	Parent	Adjust- ment	.32	.06	Parent	Adjust- ment	.28	.12
	Intro- version- extra- version	-.01	.39		Intro- version- extra- version	.12	.41		Intro- version- extra- version	.08	.33

Anxiety" was .45, and the correlation between the rating of introversion-extraversion and the second-order questionnaire analogue was .61. Comparison between the simple ratings and the second-order factor scores developed from rating schedules yielded a mean r of .56 for adjustment and .66 for introversion-extraversion.

Correlations of such magnitude show that the measures have considerable to do with each other, but are not exactly the same. The correlations do not tell which indexes are better, and any a priori assumption that the measures derived from the questionnaire and the rating schedules are the "true" measures which the simple ratings merely approximate cannot be justified. The ratings have as strong a claim to legitimacy as any other measures, however long, arduous, and expensive the development of those measures may have been. Examination of quality requires separate tests, and information on concurrent validity for the two 7-point ratings is given in Table 8.

All of the correlations for theoretically related dimensions are statistically significant. None of the others are. The mean correlation for ratings which ought to match is .33. The mean for hypothetically unrelated ratings is .08. This is the widest discrepancy between such values which

has appeared so far. According to criteria of convergent and discriminant validity, the two ratings therefore appear to be equal or superior to the considerably more lengthy and cumbersome questionnaire, and the more complex and cumbersome rating schedules examined in this review. When considerations of cost are introduced (Cronbach & Gleser, 1957), the two ratings are vastly superior to the other devices.

From the usual dogma about test length and reliability one would expect the reliability of single items to approach zero, but these do not. Stability correlations, over a 5-week span, with an examination and a vacation between the two ratings, were .61 for adjustment and .73 for extraversion-introversion. That is high enough for use in many kinds of research, and perhaps that is all any existing "personality tests" are good for.

Reducing verbal description of the complex and mysterious human personality to two ratings may seem a trifle overdone, but it is strategically preferable to begin with a minimum of constructs, measured as efficiently as possible, and elaborate as necessary, than to start with a complex structural system, measured laboriously, and proceed in the direction of still further complexity.

DISCUSSION

For anyone interested in developing a trait theory of personality, correlations of .30 are no cause for rejoicing. Self-reports and the reports of others show very limited correspondence. The reports of different sets of judges have little to do with each other. The unities therefore appear to reflect something other than covariation among observed behavior tendencies.

Among the many possible interpretative alternatives, one of the more plausible is provided by the work of Osgood, Suci, and Tannenbaum (1957, 1962) on the measurement of meaning. When meaning structures are assessed through use of semantic differential ratings, three factors consistently recur: (a) an evaluative dimension which refers to the "goodness-badness" of objects, (b) an activity dimension, i.e., "activity-passivity," and (c) a potency dimension, i.e., "strength-weakness." It does not seem to matter who assigns the meanings. The factors have emerged in a number of different societies, with raters who vary in many ways, including degree of sophistication about the objects perceived. Particular structures may vary somewhat with changes in the pool of concepts, but the generality of the meaning dimensions over people seems very well documented.

The invariant "personality" dimensions discussed above are rather easily construed as topical variants of more general ways of attributing meaning to objects, in this case human objects. "Adjustment" is good, "neuroticism" is bad. Extraversion "means" strong and active. When judges evaluate parental attitudes (Schaefer, 1961), the most strikingly consistent factors are affection (good) and control (strong and active). When variance along

the first evaluative dimension is attenuated by including only "bad" items, such as problems, the resulting factors may reflect only the attribution, to self or others, of "dynamism," i.e., activity and strength, or passivity and weakness. Thus intercorrelations among clinical problems reduce to "conduct problems" and "personality problems" dimensions (Peterson, 1960). And factors among questionnaire items which discriminate between delinquent and nondelinquent boys reflect activity-strength and passivity-weakness in the topical form Quay and Peterson (Peterson, Quay, and Cameron, 1959; Peterson, Quay, and Tiffany, 1961) have called "psychopathy" and "neuroticism."

None of the evidence in this review requires the abandonment of trait theory, though the author believes it does suggest a considerable modification of the usual forms of trait theory. When one deals with verbal judgments, the behavior of subjects, the perceptions of observers, and the properties of situations can all contribute to patterns of variation and covariation among the ratings, and all must be included in any comprehensive, operationally defined descriptive system. Relative strengths of influence can only be distinguished by systematic variation of subject behavior, observer tendencies, and situational characteristics, with at least two independent procedures for examining each. Designs of adequate complexity are seldom employed, and until such designs are used more systematically most of the important questions about situation, role, personality, and behavior will remain unanswered.

IMPLICATIONS

If broad group factors are generally more efficient and stable than the narrow "primary" factors with which

most investigators deal, attention should probably be concentrated on the former. At least higher order factors should not be omitted from research, as is so often done when conventional criteria of factor inclusion are employed. Programs of descriptive research should begin with the smallest possible number of constructs. The system should be expanded and refined only as new information may require.

If one wishes to deal with more specific aspects of personality, conventional factor analysis seems to offer little advantage as a data-reducing operation. To date, most narrow primary factors in the personality realm have not convincingly passed the tests of efficiency and invariance. If this is so, and if factors can be fractionated infinitesimally, psychology's descriptive dilemma will be unresolved by continuing to generate more and more dimensions with each new analysis. In any single investigation, study of more numerous and limited behavior segments can usually proceed as well by examination of the elemental variables themselves.

If the major factors which emerge from analysis of verbal data are generally more relevant to idiosyncratic meaning structures than to behavioral traits, the scope of any inference drawn from verbal responses to verbal stimuli should be sharply limited at the start, and extended only as new relationships permit. At the strongest, one should speak of attitudes toward the self or other objects. Such an inference is not devoid of significance, but it has some bounds.

If a "personality" factor fails, as most seem to have failed, to exhibit satisfactory generality over media and across situations, inference to "trait" or any analogous personality construct is likewise unjustified. One is forced to use qualifiers, e.g., "adjustment as

seen by teachers in a school setting." This is cumbersome, but it may be necessary. Contemporary assessment practices emphasize extensive testing of the individual in a single setting. If people are as "different" as they appear to be in different situations, and when viewed from different perspectives, it is clearly essential to seek more information than most psychologists presently do about these situations and from those perspectives.

If it is possible to get the critical information in most ratings and questionnaires with drastically abbreviated instruments, it may be possible to obtain more useful information in other ways at no great increase in assessment time.

REFERENCES

- BECKER, W. C. The matching of behavior rating and questionnaire personality factors. *Psychological Bulletin*, 1960, 57, 201-212.
- BURT, C. Group factor analysis. *British Journal of Psychology, Statistical Section*, 1950, 3, 40-75.
- CAMPBELL, D. T., & FISKE, D. W. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 1959, 56, 81-105.
- CATTELL, R. B. Confirmation and clarification of primary personality traits. *Psychometrika*, 1947, 72, 402-421.
- CATTELL, R. B. Factor rotation for proportional profiles: Analytical solution and an example. *British Journal of Psychology, Statistical Section*, 1955, 8, 83-92.
- CATTELL, R. B. *Personality and motivation structure and measurement*. New York: World, 1957. (a)
- CATTELL, R. B. *Handbook for the 16 personality factor questionnaire*. Urbana, Ill.: Institute for Personality and Ability Testing, 1957. (b)
- CATTELL, R. B. Extracting the correct number of factors in factor analysis. *Educational and Psychological Measurement*, 1958, 18, 791-838.
- CATTELL, R. B. Theory of situational, instrument, second order, and refraction factors in personality structure research. *Psychological Bulletin*, 1961, 58, 160-174.

- CATTELL, R. B. The basis of recognition and interpretation of factors. *Educational and Psychological Measurement*, 1962, 22, 667-698.
- CATTELL, R. B., & BAGGALEY, A. R. The salient variable similarity index for factor matching. *British Journal of Statistical Psychology*, 1960, 13, 33-46.
- CATTELL, R. B., & COAN, R. W. Child personality structure as revealed by teachers' behavior ratings. *Journal of Clinical Psychology*, 1957, 13, 315-327.
- CATTELL, R. B., & GRUEN, W. The personality factor structure of 11-year-old children in terms of behavior rating data. *Journal of Clinical Psychology*, 1953, 9, 256-266.
- CRONBACH, L. J., & GLESER, GOLDINE C. *Psychological tests and personnel decisions*. Urbana, Ill.: Univer. Illinois Press, 1957.
- HARMAN, H. H. *Modern factor analysis*. Chicago: Univer. Chicago Press, 1960.
- HUMPHREYS, L. G. Note on the multitrait-multimethod matrix. *Psychological Bulletin*, 1960, 57, 86-88.
- HUMPHREYS, L. G. The organization of human abilities. *American Psychologist*, 1962, 17, 475-483.
- LINN, R. L. Use of random normal deviates to determine the number of factors to extract in factor analysis. Unpublished master's thesis, University of Illinois, 1964.
- MEELAND, T. An investigation of hypotheses for distinguishing personality factors A, F, and H. Unpublished doctoral dissertation, University of Illinois, 1952.
- OSGOOD, C. E. Studies on the generality of affective meaning systems. *American Psychologist*, 1962, 17, 10-28.
- OSGOOD, C. E., SUCI, G. J., & TANNENBAUM, P. H. *The measurement of meaning*. Urbana, Ill.: Univer. Illinois Press, 1957.
- PETERSON, D. R. The age generality of factors derived from ratings. *Educational and Psychological Measurement*, 1960, 20, 461-474.
- PETERSON, D. R. Behavior problems of middle childhood. *Journal of Consulting Psychology*, 1961, 25, 205-209.
- PETERSON, D. R., & CATTELL, R. B. Personality factors in nursery school children as derived from teacher ratings. *Journal of Consulting Psychology*, 1959, 23, 562.
- PETERSON, D. R., QUAY, H. C., & CAMERON, G. R. Personality and background factors in juvenile delinquency as inferred from questionnaire responses. *Journal of Consulting Psychology*, 1959, 23, 395-399.
- PETERSON, D. R., QUAY, H. C., & TIFFANY, T. L. Personality factors related to juvenile delinquency. *Child Development*, 1961, 32, 355-372.
- SCHAEFER, E. S. Converging conceptual models for maternal behavior and for child behavior. In J. C. Glidewell (Ed.), *Parental attitudes and child behavior*. New York: Charles C Thomas, 1961.
- TUCKER, L. R. A method for synthesis of factor analysis studies. Personnel Research Section, Department of the Army, Report No. 984, 1951.
- VERNON, P. E. *The structure of human abilities*. New York: Wiley, 1950.
- WETZEL, L. C. Personality traits and cognitive meanings. Unpublished master's thesis, University of Illinois, 1963.

(Received December 6, 1963)

THE EBB OF RETENTION¹

HARRY P. BAHRICK

Ohio Wesleyan University

Retention curves based upon conventional techniques of representation are shown to be of limited usefulness because they confound the measurement of decay of memory traces with the effects of several other variables. The additional variables arise from the relations among the threshold of recall or recognition, and the mean, variance, and shape of the distribution of associative strength. A new method of representing the retention process is proposed. According to this method, estimates of changes in the position of the mean of the distribution of associative strength are plotted in units of variance called "ebbs." The assumptions underlying the use of ebbs are examined and evidence regarding the validity of the assumptions is presented.

CRITIQUE OF PRESENT METHODS OF REPRESENTING RETENTION

It is convenient to differentiate two types of indicants of retention for the purpose of the present discussion. The first type is based on a dichotomy of performance; the second refers to savings scores based upon relearning of the material. The term dichotomous means that in ordinary use the indicant yields only two scoring categories for the observed response, i.e., the subject (*S*) either recalls correctly, or he fails to recall; he either recognizes correctly, or he does not. It is true that additional scoring categories are possible, e.g., classification of types of errors made, or of reaction times, but these have not formed the basis for most retention curves shown in the literature. Usually authors report the percentage of correct responses on the ordinate, and time on the abscissa. The interpretation of these curves presents problems which are quite distinct from those associated with the interpretation of curves based upon savings scores

(the only nondichotomous indicants), and for this reason the two types of indicants will be discussed separately.

INTERPRETATION OF RETENTION CURVES BASED UPON DICHOTOMOUS INDICANTS

In typical verbal learning and retention studies *Ss* are repeatedly presented a sequence of verbal items, or a number of paired associates. Tests are later made to see how many of the items are recalled or recognized correctly, and if these tests are made at several points in time a retention curve can be drawn to trace the change of performance. A negatively accelerated, or classical curve, first reported by Ebbinghaus in 1885, is frequently observed when recall measures or savings scores are used. In a previous article (Bahrick, 1964) it was pointed out that past interpretations of the slope of such curves are equivocal because the slope is influenced by artifacts of the distribution of associative strength. These arguments will now be reviewed and elaborated.

The assumption of a continuous distribution of associative strength is basic to the argument. This simply means that when several *Ss* learn a list of

¹ This research was supported in whole by the United States Public Health Service, Grant MH-05685-02 to Ohio Wesleyan University. The writer wishes to thank David Bakan for valuable criticisms.

paired associates or a sequence of items, the associations they have formed at the end of training will probably not all be of exactly the same strength. Rather, if associative strength can vary along a continuum, some associations will be stronger than others, and this will be true for the several associations formed by each *S*, as well as among *Ss*, and one can thus conceive of an overall distribution of associative strength. The measurement of retention by a particular dichotomous indicant can tell us no more about this distribution than the percentage of associations which are above and below the momentary threshold value related to that indicant. Thus if 10 *Ss* have practiced a list of 10 paired associates, we might find that of the total of 100 responses required on a recall test 40 are correct and 60 are wrong. The continuity assumption demands that some of the 40 correct responses will reflect associative strength barely above the threshold of recall, while others will be far above threshold, and some of the 60 wrong responses will be barely below threshold strength, while others are far below threshold strength. The dichotomous recall score fails to reveal such additional information about the distribution. If a later recall test yields a lower retention score, i.e., fewer correct responses, the comparison tells us the number or percentage of associations that have passed below the threshold value during the time interval in question. It will now be shown that this percentage drop in a retention curve depends upon (a) the variance of the distribution of associative strength, (b) the relation of the threshold of the indicant to the mean strength of the distribution, (c) the shape of the distribution, and (d) the actual rate at which the associations weaken in time. These several factors are inextricably

confounded in the retention curve because no independent measures are available for them, and thus the customary interpretations of the slope of the curve as a reflection of the rate of decay of memory traces becomes equivocal and is frequently misleading.

Slope as a Function of Variance of the Distribution

The effect of variance of the distribution upon the slope of the curve can best be illustrated by conceiving of an extreme case. Let us assume that all associations formed by all *Ss* are exactly of the same strength, i.e., the variance of the distribution is zero; and let us further assume that all associations weaken at the same rate in time, so that the variance remains zero during the course of forgetting. It would follow that any retention curve based upon dichotomous indicants would show 100% correct performance until the associations had been weakened sufficiently to reach the threshold value. At this point they would all simultaneously be "forgotten" and the curve would show a step-function drop from 100% to 0% of retention, regardless of the rate at which the associations had weakened. The rate of weakening and the original strength in relation to the particular threshold used would determine only the time at which the drop from 100% to 0% would occur, but not the slope itself. If the variance is larger than zero, i.e., if some associations are stronger than others, the weaker ones would cross the threshold before the stronger ones, yielding something other than a step-function drop. Other things equal, the larger the variance of associative strength, the greater the time intervals separating the threshold crossing of successive associations, and the more gentle the slope of the retention curve between

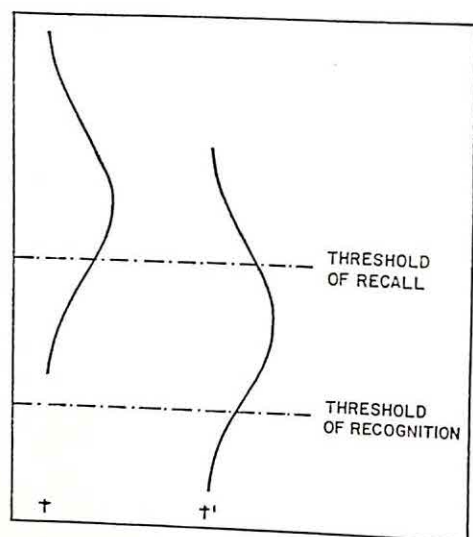


FIG. 1. A normal distribution moving past two thresholds.

two points in time. This generalization would hold regardless of the mean rate at which associations lost their strength.

Slope as a Function of Mean Associative Strength in Relation to the Indicant

It is apparent that the thresholds related to several dichotomous indicants need not be the same. Thus, it is generally thought that the threshold for recall is higher than for recognition, i.e., that the association S has formed may be sufficiently strong for a correct recognition response, but not for a correct recall response. (It was demonstrated in another paper by Bahrack and Bahrack [1965] that this is not an intrinsic difference, but depends upon the choices offered on the recognition test.) One means of conceptualizing this in relation to the distribution of associative strength is shown in Figure 1. At Time t the distribution is partly below the recall threshold, but completely above the recognition threshold. Thus, performance on recall might be 80% correct, and on recognition 100%. At Time t' the distribution has lost strength, and

has fallen for the most part below the recall threshold, but has only begun to fall below the recognition threshold. The new scores might be 20% correct for recall and 90% for recognition. Thus, during the interval from t to t' the slope of the curve based upon the recall indicant is much steeper than the slope of the curve based upon the recognition indicant. Recall performance has dropped by 60%, recognition performance by only 10%, despite the fact that the rate of weakening of the associations has been assumed constant. Thus, the sensitivity of a particular indicant depends upon the momentary threshold of the indicant in relation to the mean of the distribution of associative strength at the outset of the time interval in question. Failure to consider such artifacts has led to overgeneralizations and misinterpretations regarding the rate of forgetting for "recognition versus recall memory" (Bahrack, 1964), as well as misinterpretations regarding the slope of learning curves (Bahrack, Fitts, & Briggs, 1957).

Slope as a Function of the Shape of the Distribution of Associative Strength

Figure 2 shows several possible distributions of associative strength gradually passing through a threshold, and below each type of distribution the related retention curves. Assuming only for the purpose of illustration that associations weaken at a constant rate in time, the considerable influence of the shape of the distribution upon the slope of the retention curve is apparent and does not require elaboration. Classical, i.e., negatively accelerated curves of retention can be obtained as the result of a distribution like the one shown in B of Figure 2, although every one of the associations lost strength at a constant rate in time.

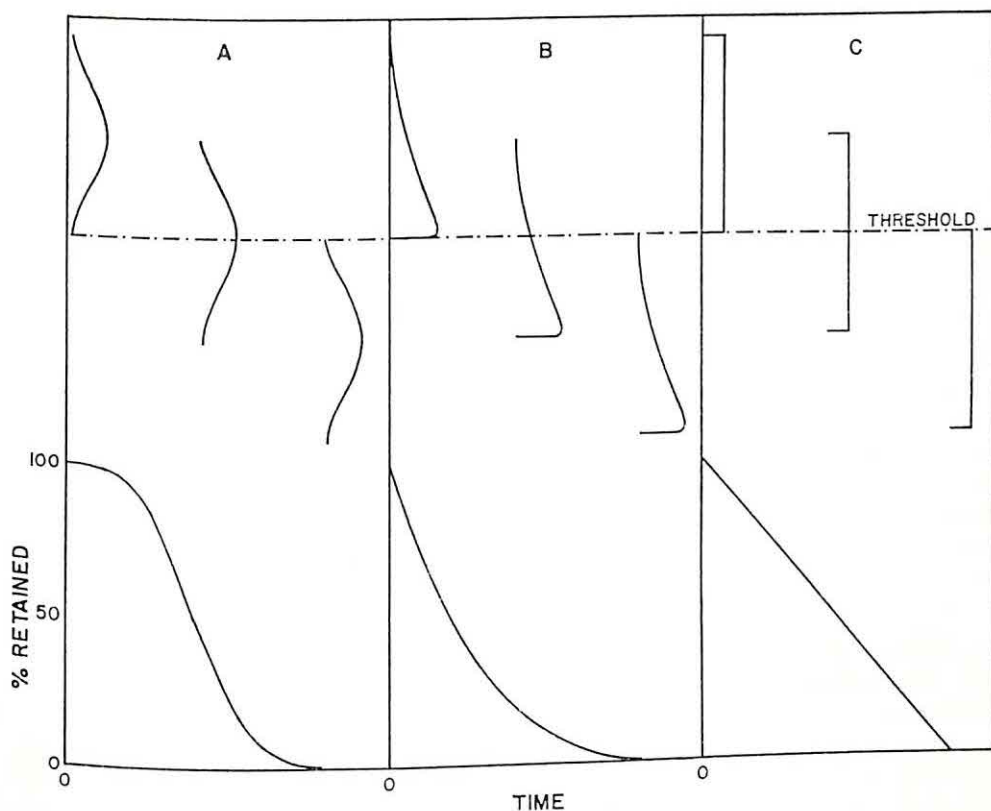


FIG. 2. The effect of the shape of the distribution of associative strength upon retention curves.

The above illustrations have shown that the percentage drop of a retention curve during a particular time interval is not only a function of the rate at which associations have weakened during this interval, but also of the variance and shape of the distribution of associative strength, and of the relation of the mean of that distribution to the threshold of the indicant being used. Failure to control or assess these various factors independently precludes meaningful comparisons of relative amounts of retention for different time intervals, or for different indicants during the same time intervals.

LIMITATIONS OF RETENTION CURVES BASED UPON SAVINGS SCORES

The only commonly used indicant of retention which is nondichotomous, i.e.,

usually yields more than two scoring categories, is based upon relearning scores. Ultimately, of course, savings scores are also based upon dichotomous measurement, i.e., the responses of Ss are still only scored as above or below the threshold of recall or recognition on each trial. However, the derived scores based upon the comparison of learning and relearning trials yield a continuum of "savings." Present criticism relates to problems which arise from (a) unreliability, and (b) the use of trials as units of measurement.

Unreliability of Savings Scores

The limitations of this method were recognized by Luh (1922) who concluded empirically that the method "produces less satisfactory results than does any other method [p. 35]." This

is true because the reliability of the comparison of learning and relearning trials is limited not only by the momentary fluctuations of dichotomous thresholds, as are all the other methods, but also by the cumulative unreliability contributed by many other variables which affect the rate of acquisition in the two learning processes to be compared.

Further, the comparison of learning and relearning rates has traditionally been limited to the point at which 100% of the material has been mastered, i.e., the point at which the entire distribution of associations has passed above a given threshold during learning. A single point of this sort does not represent a very reliable base for comparison, and the 100% point is less satisfactory than other points on the distribution, e.g., the 50% point. A more reliable procedure would base the savings score upon an average comparison of several points, e.g., the savings effected in mastering 25%, 50%, and 75% of the material. Similar principles were recognized long ago in connection with psychophysical problems, but they have not been applied to the methodology of retention studies where they are equally relevant.

Use of Trials as Units of Measurement

If retention is determined by a comparison of the number of learning trials with the number of relearning trials, it follows that trials have become the units of retention. The use of such units can form the basis for inferences about changes in associative strength only if we can assume that the gain in associative strength, or the amount learned per trial, is reasonably constant. If we assume, however, that the learning process is negatively accelerated, i.e., that progressively less associative strength is developed on

successive trials, then saving the last trials has a very different meaning from saving the earlier trials. A negatively accelerated curve of forgetting could result, even though associations weakened at a constant rate in time. This could occur because many of the last trials during which little associative strength was gained are saved at first, and later progressively fewer of the earlier trials are saved, during which progressively more was learned.

The loss of associative strength reflected by a single trial (i.e., the size of the unit) also varies as a function of individual differences in rate of learning. Thus, if an association has barely fallen below threshold and is relearned in a single trial, this small amount of forgetting would be represented as a saving of 80% for an individual who originally learned in five trials, but the same loss of associative strength would be reflected by a saving of 90% for another *S* who originally learned in 10 trials. It is apparent that the slope of retention curves based upon savings scores cannot be interpreted in terms of losses of associative strength in other than an ordinal sense.

USE OF VARIANCE UNITS OR EBBs

If retention curves are to yield information about relative rates of forgetting, i.e., the rate of weakening of associations during different time periods, it is necessary to use equal intervals on the ordinate. Further, we must either experimentally control the variables causing distributional artifacts, or develop techniques which are less sensitive to them. One possible technique is proposed here. It involves plotting estimates of the change of the mean of the distribution of associative strength in units of variance of the distribution on the ordinate. The original observations can be based upon

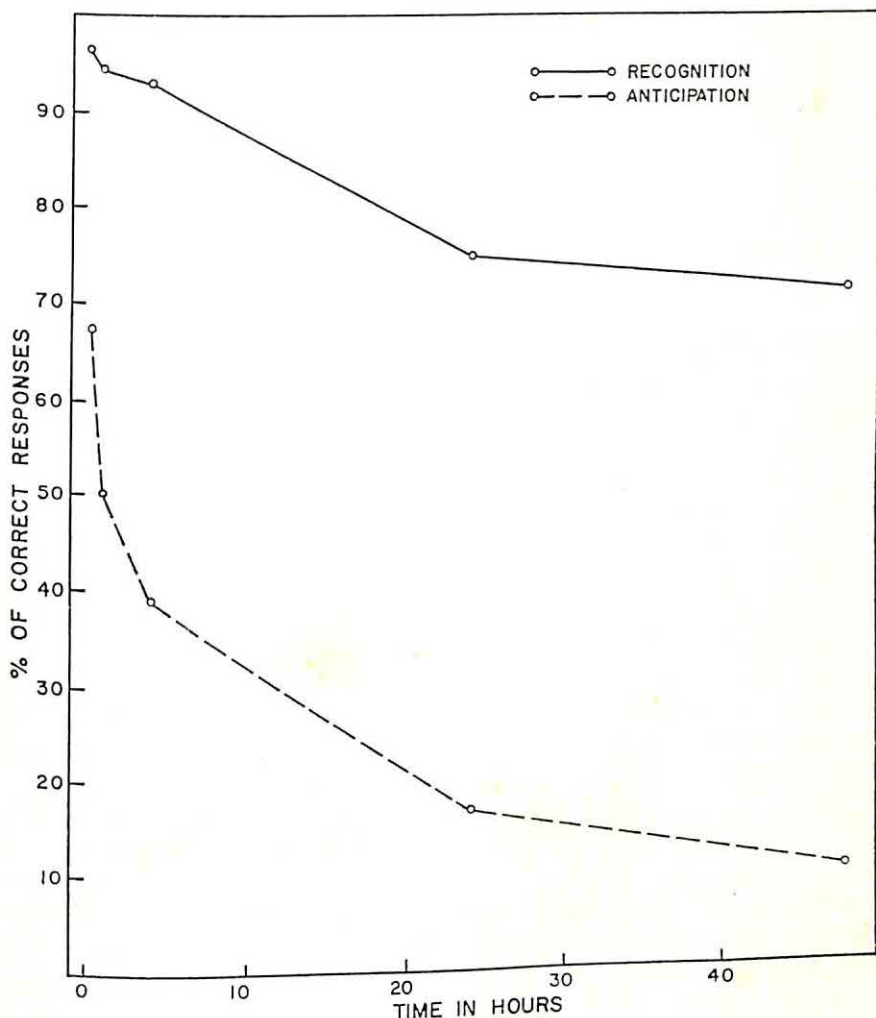


FIG. 3. Percentage correct anticipation and recognition from Luh (1922).

recall, recognition, or any combination of indicants referring to different thresholds, but the thresholds are used only as reference points for locating the mean of the distribution.

To illustrate the technique, the curve for anticipation performance in Luh's study (1922) which is plotted by the conventional method in Figure 3 can be compared with the same curve replotted in ebb units in Figure 4. Each observed percentage anticipation value is converted into a Z value by consulting a table of integrals of the normal curve. The Z value gives the distance

from the threshold point to the mean of the distribution in sigma units. If more than 50% of the responses are initially correct, as is the case here, the mean of the distribution is above the threshold and the first Z value is positive. As forgetting progresses to the point where only 50% is recalled the mean of the distribution has reached the threshold, and the Z value has become zero. When the percentage correctly recalled is less than 50 the mean of the distribution has passed below the threshold, and the Z value becomes negative. To designate

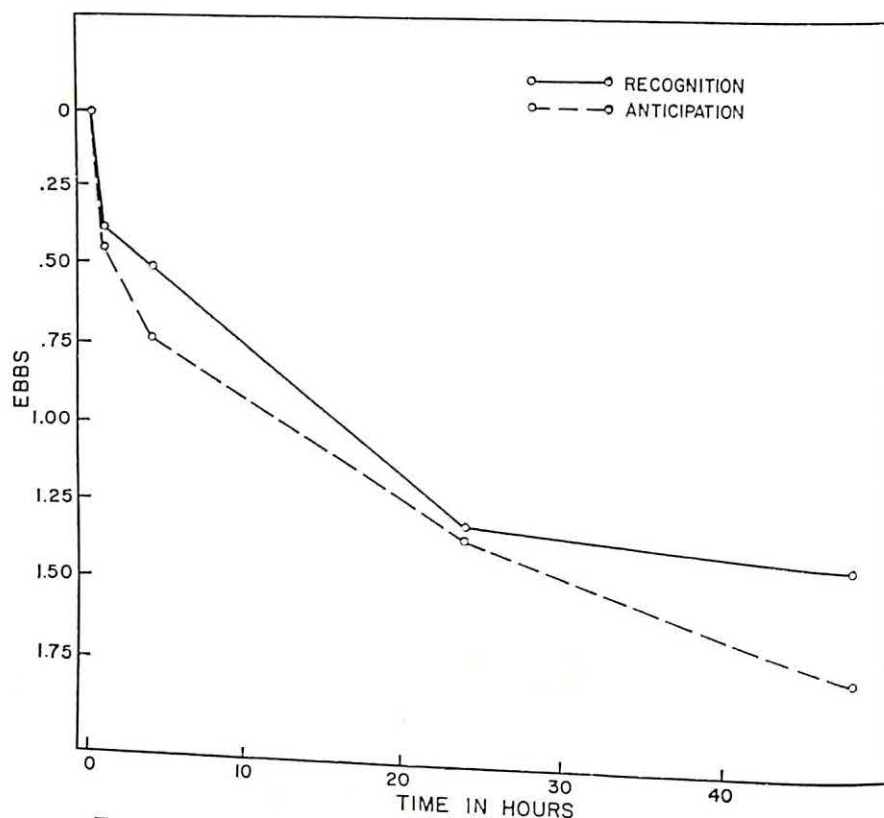


FIG. 4. Retention curves based on ebbs from data by Luh (1922).

the beginning of the forgetting process an ordinate value of zero may arbitrarily be assigned to the first Z score after learning has terminated. Subsequent values on the ordinate are obtained by subtracting each converted Z score from the original one, and thus the course of forgetting is plotted in sigma units or "ebbs," reflecting the

downward movement of the mean of the distribution of associative strength. The percentage anticipation, Z conversion, and ebb values are shown in Table 1.

ADVANTAGES OF THE METHOD

Conventional retention curves based upon dichotomous scoring indicate the percentage of associations which pass below a threshold value during a given time interval. It has been shown why the slope of such curves cannot provide a direct basis for comparison of losses of associative strength during different time intervals. The transformation proposed here does permit such a comparison, provided certain assumptions can be met. The curve based upon ebbs shows how much the mean association has weakened. The use of

TABLE 1
PERCENTAGE ANTICIPATION FROM LUH
CONVERTED TO EBBS

Percentage anticipation	Z transformation	Ebbs
67.8	.46	0
50.2	.06	.40
39.0	-.28	.74
17.8	-.92	1.38
10.0	-1.29	1.75

sigma units on the ordinate prevents the slope of the curve from being a function of the variance of the distribution, as long as the variance remains constant. This is true for the same reason that conversion of raw scores to Z scores achieves comparability despite differences in mean and sigma of the raw score distribution. While we are no better off in comparing "the amount" of forgetting for two different studies, unless we can compare the variance of the distributions involved, it is possible to compare the curves yielded by different indicants over various time intervals within the same study, provided the variance has stability. By conventional methods it has not been possible to determine whether recognition, recall, and other indicants yield basically different time functions for forgetting, because the obtained curves were differentially affected by artifacts of changing sensitivity of the indicants, illustrated in Figure 1. The proposed technique permits a reexamination of such questions under more favorable conditions.

The shape or form of the distribution continues to have an effect upon the slope of the ebb curve. The rectilinear and normal distributions shown in Figures 2A and 2C will now produce identical curves, and this is equally true for any other type of symmetrical distribution, provided an unbiased estimate of the position of the mean is obtained. If no unbiased estimate can be obtained, as is likely for distributions departing markedly from normality, the obtained functions will be affected, as will be shown later on in more detail. Artifacts produced by using different indicants such as recall and recognition, illustrated in Figure 1, will disappear, if we are dealing with a normal distribution, but continue to be present if the distribution departs markedly from

normality. If the distribution is normal and maintains constant variance, and if recall and recognition refer to different thresholds dichotomizing the same distribution at different levels, then it should make no difference at all which of the thresholds we use as our reference point in tracing the course of the mean of the distribution. Departures from normality will result in different types of bias in locating the mean with respect to different thresholds, and thus, the method provides a potential source of information about the shape of the distribution. An illustration of this use of the method will be given later.

ASSUMPTIONS UNDERLYING THE USE OF THE PROPOSED TECHNIQUE

The assumptions, already implied above, are of two types. The first, and most important, refers to the variance of the distribution; the second, to the thresholds of different indicants in relation to the distribution. Plotting changes of the mean in sigma units is meaningful only if the sigma of the distribution remains relatively constant throughout the course of forgetting. It would not matter if individual associations lost strength more quickly than others, as long as compensatory variations of this type left the variance of the entire distribution stable.

Direct tests of the validity of this assumption are not possible, since no direct measures of variance are applicable. The strength of individual associations can be measured only in relation to dichotomous thresholds. Even if one accepts the number of correct responses during training as indicative of associative strength, this indicant would yield only a single measure of variance, and no information regarding the stability of the vari-

ance during the retention period could be deduced from it.

The second group of assumptions relates to the question as to whether different dichotomous indicants such as free recall, anticipation, and recognition divide the *same* distribution at different levels. Alternatively, one might assume that these indicants, though of different difficulty level, do not assess the same associations, but rather relate to different types of concurrent learning, and thus refer to separate distributions. The most relevant evidence in favor of a unitary assumption is statistical. It has been shown repeatedly (Postman, Jenkins, & Postman, 1948) that high correlations exist between items correctly recognized and those recalled, i.e., that practically all of the items correctly recalled are also correctly recognized. (But not vice versa.) If the learning processes were independent such high correlations would not occur. Accordingly, Postman, Jenkins, and Postman concluded: The basic difference between the two tests appears to lie in the minimal strength of association which they require for successful performance.

An important qualification must be stated, however. Performance on recall, as well as on recognition tests, may or may not depend upon memory of sequences. In a free-recall test, for example, *S* may be asked to write down as many of the items as he can recall, and the responses can be scored without regard to the sequence of presentation. Likewise, on a recognition test *S* may be asked to identify the items among other incorrect ones, without regard to sequence of presentation. The assumption of a single distribution dichotomized at two levels appears reasonable in this situation. It is also possible, however, to score sequential associations in either the recall or the recognition test. The method of anti-

cipation demands such association on the recall level. To test these on the recognition level a variation of the procedure of reconstruction can be employed, in which the correct items are presented in various sequences, and *S* must identify the correct sequence. It is reasonable that learning in a paired-associate task involves a two-phase process, where sequential associations are formed only after individual items have been differentiated. If one, but not both the recall and recognition tasks requires memory of sequences, the assumption of a single distribution dichotomized at two levels would seem less justified.

INDIRECT EVIDENCE RELATED TO THE ASSUMPTIONS

Investigations in which several types of indicants are used to test retention at several points in time are a potential source of evidence in regard to both types of assumptions previously discussed. If the variance of the distribution remains constant during the course of forgetting, and if this same distribution is dichotomized by two thresholds of different difficulty level, performance differences on the two tests, when converted into *Z* scores, should remain relatively stable over the total retention period. Referring to Figure 1, we would expect equal *Z*-score intervals to separate performance on recall and recognition tests at Time *t'*, and any later time at which both thresholds dichotomize the distribution. The accuracy of this type of check is limited because the relevant investigations either (a) use the same *Ss* repeatedly and thus confound the later measures or (b) use different *Ss* at successive time intervals, and for different measures, and thus fail to provide a possibility of reexamining the same distribution. In the latter case they provide only another sample of a

TABLE 2
Z SCORE DIFFERENTIALS IN SEVERAL STUDIES COMPARING RECALL AND
RECOGNITION PERFORMANCE

	Luh				
	20 minutes	1 hour	4 hours	1 day	2 days
Percentage correct recognition	97.8	94.6	93.3	74.6	71.5
Percentage correct anticipation	67.8	50.2	39.0	17.8	10.0
Differences in percentage	30.0	44.4	54.3	56.8	61.5
Recognition Z value	2.01	1.61	1.50	.66	.57
Anticipation Z value	.46	0.06	-.28	-.92	-1.28
Difference in Z values	1.55	1.55	1.78	1.58	1.85

	Miler					
	5 seconds	1 hour	6 hours	24 hours	4 days	14 days
Percentage correct recognition	65.8	65.8	54.1	55.0	50.4	38.3
Percentage correct recall	41.6	35.0	26.6	27.5	21.6	6.6
Difference in percentage	24.2	30.8	27.5	27.5	28.8	31.7
Recognition Z value	.41	.41	.10	.13	.01	-.30
Recall Z value	-.21	-.39	-.63	-.60	-.79	-1.51
Difference in Z values	.62	.80	.73	.73	.80	1.20

	Bahrick and Bahrick			
	Immedi- ately	2 hours	2 days	2 weeks
Percentage correct—easy recognition	80.0	85.7	77.1	38.6
Percentage correct—anticipation	78.6	62.9	65.7	35.7
Difference in percentage	1.4	22.6	11.4	2.9
Easy recognition Z value	.84	1.07	.74	-.29
Anticipation Z value	.79	.33	.40	-.37
Difference in Z values	.05	.74	.34	.06

	Bahrick and Bahrick			
	Immedi- ately	2 hours	2 days	2 weeks
Percentage correct—difficult recognition	54.3	28.6	41.4	14.3
Percentage correct—anticipation	78.6	62.9	65.7	35.7
Difference in percentage	24.3	34.3	24.3	21.4
Difficult recognition Z value	.11	-.57	-.22	-1.07
Anticipation Z value	.79	.33	.40	-.37
Difference in Z values	.68	.90	.62	.70

similar distribution. Another limitation arises from the type of tests used. These may require sequential learning for one, but not both of the tests, and thus they may relate to different stages of a two-stage learning process, as discussed above.

Table 2 summarizes data from three different investigations in which two or more indicants of retention were recorded at different time periods. The stability of the Z-score differentials is most impressive for the recognition and anticipation performance in Luh's study

where the same *Ss* were used repeatedly. Despite the fact that differences between recognition and anticipation performance increase from 30 to 60%, the *Z*-score differentials remain remarkably constant during the entire range of time from 15 seconds to 2 days. The data of Miler show stability of *Z*-score differentials for all except the 2-week interval, and the data by Bahrick and Bahrick (1965) show some stability for the differentials between a difficult recognition task and an anticipation task, but not for the comparison of an anticipation and an easy recognition task. A reminiscence effect in performance on the ease recognition task may account in part for the instability, as well as the assessment of sequential learning in anticipation performance, but not in connection with recognition performance. A search of the literature revealed several other studies in which two or more indicants of retention were recorded at different time intervals, but they did not lend themselves to the above type of analysis. Data from a study by Myers (1914) are not presented because they are based upon only two time intervals. Data from studies by Postman and Rau (1957) and by Burt and Dobell (1925) could not be analyzed. In the former case, recognition performance remained nearly 100% correct throughout the retention period, and thus the threshold of recognition did not dichotomize the distribution so as to yield a reliable reference value (a situation similar to the one shown in Figure 1 at Point *t*). In the latter case, recall performance remained near 0% after the first measurement. In this connection it is important to note that percentage values become less reliable as they approach the values of either 100% or 0% and that this unreliability is magnified by a *Z*-score transfor-

mation which increases differentiation of percentage values near the extremes. Consequently the transformation described here should probably not be used when scores on the dichotomous task are extreme. This would apply to any study in which training continued to the point at which 100% of responses are correct, or above threshold.

It is not possible to assert on the basis of the evidence from these studies that the variance of distributions of associative strength does not change during the course of retention. No statistical test of the significance of changes in the *Z*-score differentials is applicable because data are not available for individual *Ss*. Even if such data were available, and failed to reveal significant changes, they would not permit a proof of the null hypothesis. If one keeps in mind, however, the fact that different *Ss* are involved for successive time intervals and that the *Z*-score differentials show no consistent trend toward increasing or decreasing over intervals ranging from 15 seconds to 2 weeks, but rather show comparative stability, the assumption of constant variance appears plausible.

It would seem necessary to assume that variance values would eventually decrease as memory traces continue to weaken, but this may occur beyond the range of sensitivity of available dichotomous indicants. Only the 2-weeks value obtained from the data by Miler suggests such a phenomenon during the time range covered by these investigations.

EVIDENCE REGARDING THE SHAPE OF THE DISTRIBUTION

The analysis presented so far lends some support to the assumption of stability of variance, but offers little evidence regarding the shape of the

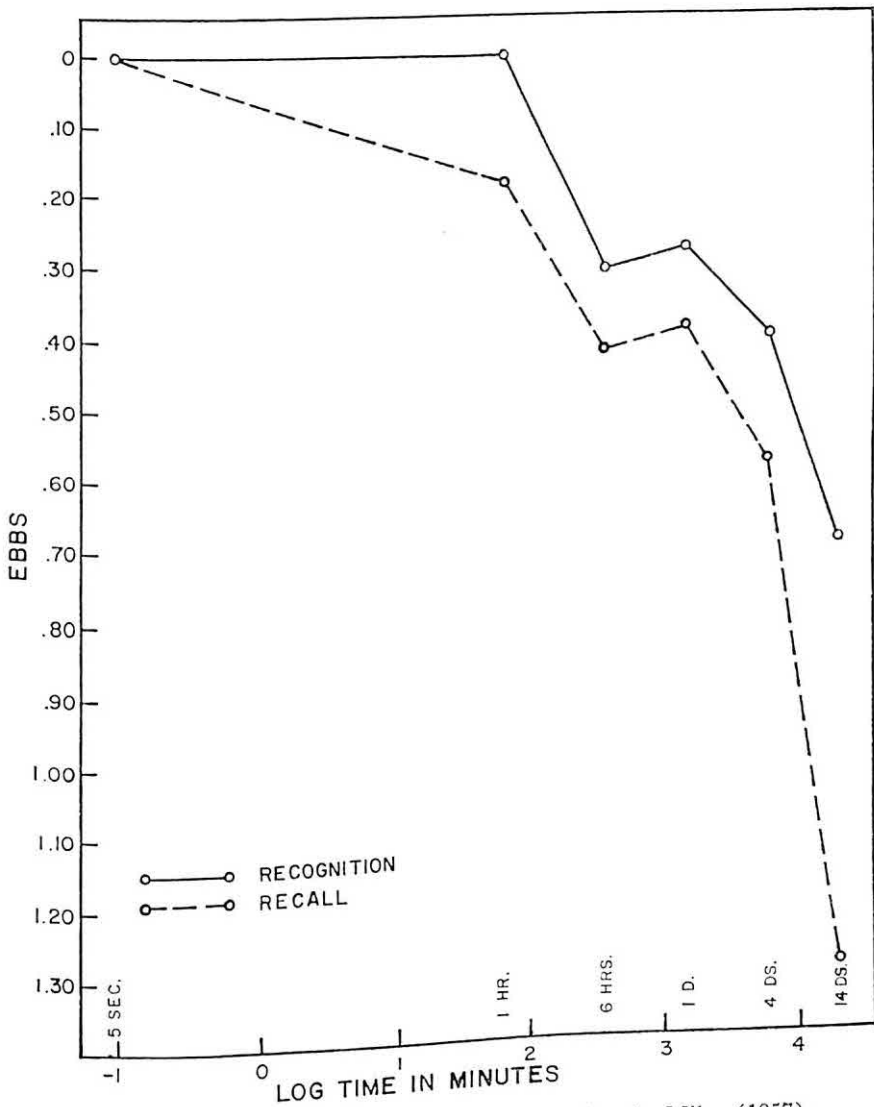


FIG. 5. Retention curves based on ebbs from data by Miler (1957).

distribution of associative strength. Figures 4, 5, and 6 show retention curves based on ebbs for the data of Luh, of Miler, and of Bahrick and Bahrick, respectively. Separate curves are derived from the recognition and recall values of each study. It must be noted that thresholds of recognition and recall are used here only as reference values, and that both curves plot estimates of the position of the mean of presumably the same distribution. The

estimates are based upon integral values of the normal distribution, and thus, divergence of the two estimates offers some clues in regard to departures of the distributions from normality.

If a distribution is less peaked than a normal one, i.e., if it has relatively more area at the extremes, we would expect exaggerated estimates of the drop in the position of the mean, when the estimate is based on percentage

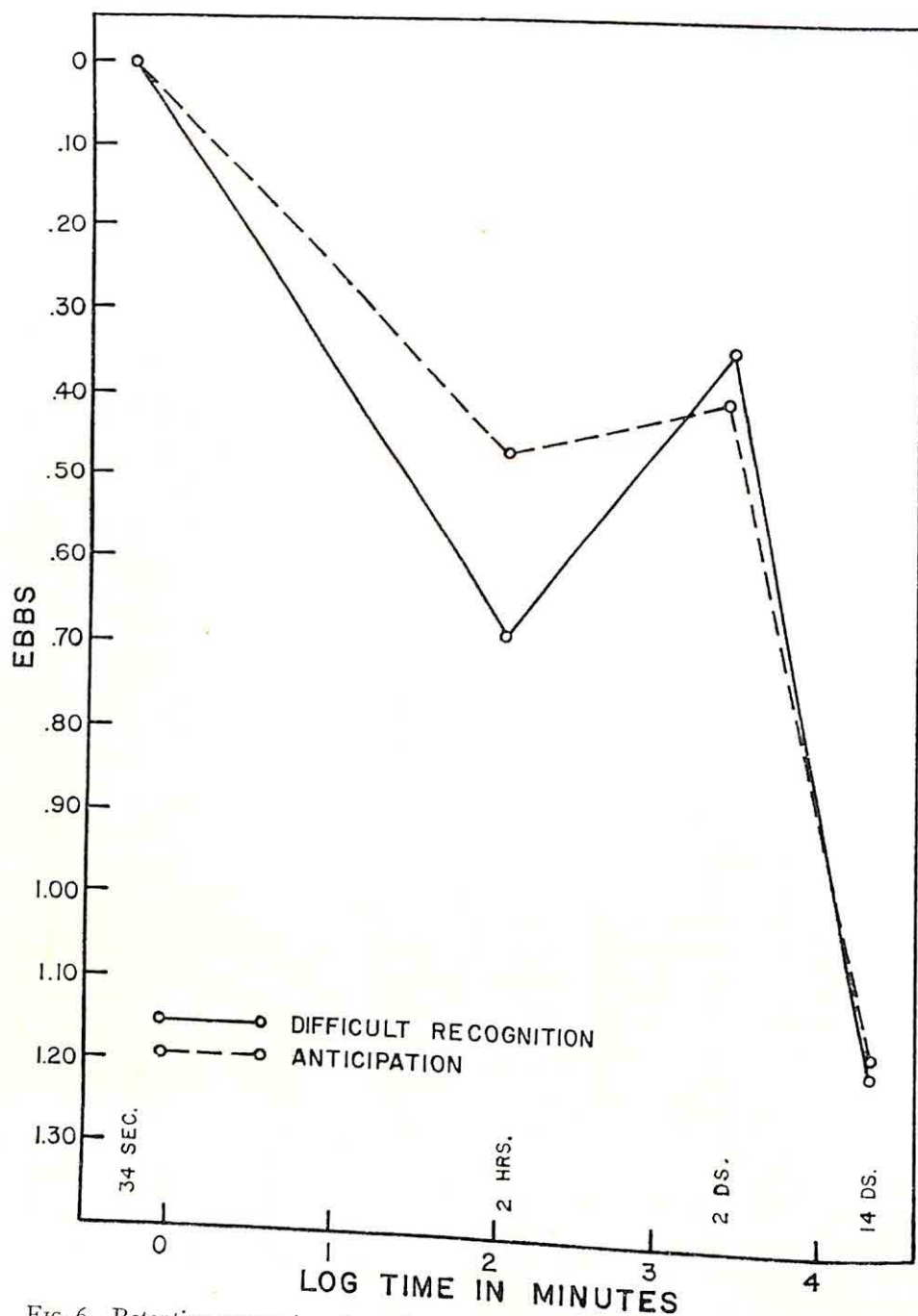


FIG. 6. Retention curves based on ebbs from data by Bahrick and Bahrick (1965).

changes recorded near the extremes of the distribution. However, if the threshold is passing near the mean of the distribution, the predicted drop in the mean would be an underestimate.

Conversely, with a distribution more peaked than a normal one, the drop of the mean would be underestimated if the reference point is at the extremes, but overestimated if the reference point

is near the center. Comparison of the magnitude of Z -score changes of the mean predicted from the two reference points (values in Table 2) shows the following: The more extreme reference point leads to smaller estimates of the drop in the position of the mean (smaller Z -score changes) when the percentage values are above the mean (larger than 50%); if the values are below the mean (smaller than 50%) the more extreme score leads to larger estimates of the drop in the position of the mean (larger Z -score changes). This is consistent for all three studies, regardless of whether the curve based upon recognition or recall is the higher one.

It would appear from these facts that the distribution is somewhat asymmetrical, i.e., that the upper part of the distribution is peaked and shows less area at the extreme than a normal distribution, but that the lower part of the distribution is skewed, i.e., shows more area in the extreme than a normal distribution.

These interpretations are tentative, and will require additional support before they can be accepted with confidence. However, it is apparent that the method which has been described offers a variety of opportunities for exploring aspects of the retention process and of the distribution of associative strength, which have not been available to the conventional techniques of representation. Further information about these distributions during the course of acquisition, as well as during the course

of retention, should prove to be of value in the analysis of phenomena of learning and retention.

REFERENCES

- BAHRICK, H. P. Retention curves: Facts or artifacts. *Psychological Bulletin*, 1964, 61, 188-194.
- BAHRICK, H. P., FITTS, P. M., & BRIGGS, G. E. Learning curves: Facts or artifacts. *Psychological Bulletin*, 1957, 54, 256-268.
- BAHRICK, H. P., & BAHRICK, P. A re-examination of the interrelations among measures of retention. *Quarterly Journal of Experimental Psychology*, 1965, 17.
- BURTT, H. E., & DOBELL, E. M. The curve of forgetting for advertising material. *Journal of Applied Psychology*, 1925, 9, 5-21.
- EBBINGHAUS, H. *Über das Gedächtnis: Untersuchungen zur experimentellen Psychologie*. (Orig. publ. 1885) [Memory: A contribution to experimental psychology.] (Trans. by H. A. Ruger & C. E. Busse-nius) New York: Teachers College, Columbia University, 1913.
- LUH, C. W. The conditions of retention. *Psychological Monographs*, 1922, 31 (3, Whole No. 142).
- MILER, ADRIENNE. Vergleich der Vergessenskurven fuer Reproduzieren und Wiedererkennen von sinnlosem Material. *Zeitschrift für experimentelle und angewandte Psychologie*, 1957, 7, 29-38.
- MYERS, G. C. A comparative study of recognition and recall. *Psychological Review*, 1914, 21, 422-456.
- POSTMAN, L., JENKINS, W. O., & POSTMAN, D. L. An experimental comparison of active recall and recognition. *American Journal of Psychology*, 1948, 61, 511-520.
- POSTMAN, L. & RAU, L. Retention as a function of the method of measurement. *University of California Publications in Psychology*, 1957, 8, 217-270.

(Received December 12, 1963)

or even reasonable, to restrict the term "defense" only to a failure to recognize a "threatening" stimulus, since an organism which fails to recognize danger would, at first glance, seem especially defenseless. It has been suggested that the term was applied in this particularly restricted way because of an analogy to the ego defenses of psychoanalytic theory (e.g., Howie, 1952). Such an analogy is false because the ego defenses are not restricted only to those which cause a failure to recognize external stimuli (e.g., Fenichel, 1945; Freud, 1937; Lazarus, 1954).

Recognition thresholds which are either raised or lowered have been interpreted as due to defenses, and this view has received experimental support (e.g., Chodorkoff & Chodorkoff, 1958; Deese, Lazarus, & Keenan, 1953; Eriksen, 1951a, 1951b, 1954a, 1954b, 1954c; Eriksen & Davids, 1955; Eriksen & Lazarus, 1952; Lazarus & Longo, 1953; Miller, 1954; Truax, 1957).

Perceptual Vigilance

Dulany (1957), among others, has argued that the really important form of perceptual defense is a "vigilant" tendency to recognize emotion-arousing stimuli unusually readily, although its presence is often obscured by response bias.

A mechanism of perceptual vigilance is both plausible and of value in theory. It need only strengthen or clarify the perceptual response during the process of recognition. It need not involve a "super-discriminating preperceiver" who selectively prevents recognition, a preperceiver who has proved difficult to find (Eriksen, 1960). A theory of vigilance has been recognized as important to the explanation of stimulus detection (Frankman & Adams, 1962).

The constructs of such a theory might well include a mechanism of perceptual vigilance.

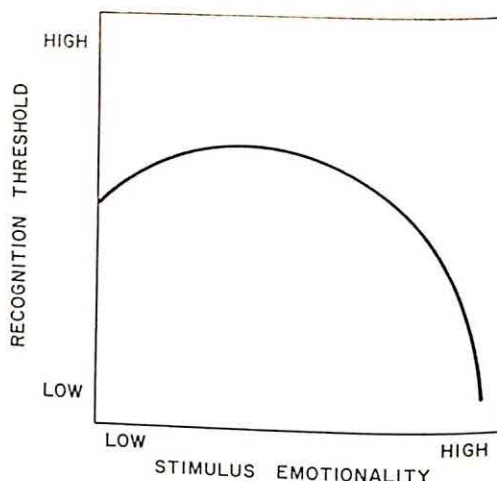


FIG. 1. The relation of recognition threshold to level of stimulus emotionality (suggested by Brown, 1961; Bruner & Postman, 1947; DeLucia & Stagner, 1953; etc.).

Formulation Integrating Previous Findings and Conceptions

In his recent monograph, Brown (1961) argues that perceptual defense is best defined as a *descriptive term for any systematic relationship between stimulus emotionality and the ease of recognition of stimuli*. He implies, as have other reviewers, that both elevated and lowered thresholds are part of the same functional relationship (Inglis, 1961).

Some of the earlier studies (e.g., Bruner & Postman, 1947; DeLucia & Stagner, 1953), recent reviews (e.g., Brown, 1961; Inglis, 1961), and an additional study by Brown (1961) support the interpretation that threshold first rises, then falls with an increase in stimulus emotionality (Fig. 1). Such a complex relationship does not support a restricted definition of perceptual defense, or the "one-tailed" statistical tests which have so frequently been criticized (e.g., Eysenck, 1961; Hick, 1952).

Criteria for Studies Evaluating a Response-Bias Hypothesis

Thus far this paper has attempted to demonstrate that studies supporting a response-bias hypothesis have often been particularly restricted in their approach. They may constitute an attack on studies using an overly restricted conception of perceptual defense. However, they do not necessarily conflict with perceptual-defense studies using the more complex conception suggested by research done since the early work (e.g., Brown covers well over 300 papers). Perceptual defense can be conveniently reviewed in the context of the criteria research has suggested for a study evaluating the response-bias hypothesis.

1. As was shown by the controversy following the McGinnies study (1949), measurement of perceptual defense requires that the experimenter avoid using unexpected or socially unacceptable stimuli. Otherwise either response suppression or a lack of familiarity with certain stimuli provide a ready explanation of results (Freeman, 1954, 1955; Howes & Solomon, 1950, 1951; Lacey, Lewinger, & Adamson, 1953).

2. It seems important to use stimuli chosen to be personally emotion arousing for the subjects for whom they are used (Eriksen, 1954a) and to incorporate adequate controls for word frequency. A desirable control for *individual* response probability would be allowed by response bias measures obtained during the course of a perceptual defense study.

3. Results should not be averaged for stimuli which differ greatly in their emotion-arousing properties if Figure 1 has any validity. The low threshold for some stimuli and the high thresholds for others would be averaged out when combined.

4. Averages computed for subjects with quite different personalities may also obscure results. The subject's personality may serve to alter the effective level of stimulus emotionality (Brown, 1961; Inglis, 1961). Brown, as he predicted, found evidence that introverts reach their highest recognition threshold at a lower stimulus emotionality than extroverts. Variables were defined by scores on the Maudsley Personality Inventory (Eysenck & Claridge, 1962).

Other individual differences in perceptual defense have been reported as well (Eriksen, 1951a, 1951b, 1952a, 1952b; Eriksen & Davids, 1955; Eriksen & Lazarus, 1952; Lazarus & Longo, 1953; Truax, 1957; Wall & Guthrie, 1959). Mathews and Wertheimer (1958) found perceptual defense significantly greater than response bias in subjects high on the MMPI *Hy* scale (Dahlstrom & Welsh, 1960) but not for those high on the *Pt* scale. It may be that some learning histories and personalities lead to *perceptual* and some to *nonperceptual* forms of defense.

5. Averages and comparisons which involve different *situations* may also serve to obscure or confuse results. The experimental situation (only partly described by instructions to subjects) may also serve to alter the effective level of stimulus emotionality. For example, Brown found that with females he was able to demonstrate the whole curve shown in Figure 1; with males, only the first (ascending) portion of the curve was demonstrated. The sexes differed by about three standard deviations in their thresholds to the most emotional words. He suggested that being alone with a male experimenter was more emotion arousing and the effective level of stimulus emotionality was greater for females than for males. *Perceptual* interpretations of modifications of perceptual de-

fense by slight changes in instructions have also been given by Blum (1954), Clapp (1951), and Nelson (1955).

The changes produced by different experimental situations and instructions have also been taken to imply that perceptual defense is *not perceptual* (e.g., Postman, Bronson, & Gropper, 1953; Nothman, 1962). Certainly studies of verbal behavior have shown the effects of the experimenter's behavior on response probabilities (e.g., Krasner, 1958; Salzinger, 1959). The subtle cues he provides either intentionally or accidentally may have complex effects (Sarason & Minard, 1963). The precise effect of instructions may depend on the nature of the task the subject performs (e.g., Sarason & Minard, 1962).

Either a wholly perceptual or a wholly nonperceptual interpretation of changes in perceptual defense seems premature for studies in which perceptual measures are thought to be affected by response bias and response bias is not measured.

6. Since there are many complex differences between subjects and situations it seems important to obtain measures of perceptual defense and response bias in the same subjects and nearly simultaneously.

AN EVALUATION OF THE RESPONSE-BIAS HYPOTHESIS

Studies have demonstrated that, under certain circumstances, response bias may be great when compared with perceptual-defense results (e.g., Goldstein, 1962). However, those meeting criteria suggested above demonstrate that response bias does not satisfactorily account for perceptual defense; it is necessary to develop explanations which allow the stimulus a greater role and which do not depend on mechanisms which have their main effects on purely verbal responses.

Mathews and Wertheimer (1958), using visually presented words, found perceptual defense over and above any evidence for simultaneously measured response suppression. Minard (1963) replicated and extended their research (which used a stimulus selection technique similar to that used by Brown in selecting highly emotion-arousing stimuli). This study will be described in some detail because it illustrates major points of this review and allows more effective evaluation of previous work.

Procedure

Subjects. A sample similar to that of Mathews and Wertheimer (1958) was selected. None who were selected refused to participate as paid volunteers uninformed regarding the basis of selection or the nature of the study. The 26 males and 26 females were college students with scores in the upper quartile of a group of 312 MMPI *Hy* scores, but below the median on the *Pt* scale. All were introductory psychology students naive regarding perceptual defense.

Stimulus Selection. Stimuli were eight words selected by objective criteria applied to a word association test consisting of 46 potentially emotion-arousing words and of neutral words paired with them on the basis of structure and frequency in English. Words for the test were selected by the agreement of three psychologists familiar with the college population. All words were common and acceptable in student conversation (e.g., "adjust," "child," "dance"). Four emotion-arousing words were chosen for each subject, and each was paired with a neutral word. Chosen stimuli had associative reaction times within the longest 8% of the subject's times, a time long also when compared with the mean time for that word as found for a norm group of the selected, like-sexed

subjects. Other evidences of emotional arousal were observed or tape recorded (blushing, sweating, stuttering, long explanations of associations, etc.). This evidence was ignored in the selection of emotional words, since it was not found to change the stimuli selected.

Stimulus Presentation. The relatively brief word association test was followed by a selection of the conditions for tachistoscopic presentation of stimuli, i.e., the settings at which the subject's discrimination of structurally similar stimuli was 60-70% accurate and better than chance according to a binomial test. In this determination the subject held a card listing two neutral stimuli not selected for later presentation (e.g., "line," "land"). He was instructed to name the word presented on each trial and to describe "exactly what he saw" on the screen. Immediately before the perceptual defense experiment the determination was repeated. Stimuli were presented with the first letter at center field by a stroboscopically calibrated tachistoscope, accurate within 5%. They subtended an average visual angle of 9.5 degrees.

After this determination each subject was given a card listing the eight stimuli selected for him, told "We'll do the same thing with different stimuli" and given the instructions used by Mathews and Wertheimer (1958) "for every flash, tell me what word you think was presented." In addition, the subject continued describing "exactly what he saw" when he wished. Stimuli were then repeatedly presented at 10-second intervals except for three regularly spaced 30-second breaks. Within each of the four blocks of 10 trials, the stimuli used were randomly arranged. Unknown to the subject, half of the eight stimuli were *never* presented. Only two emotional words and their neutral pairs had been selected for presentation

according to a procedure which equated presented and un-presented words for emotionality on the basis of all available information. In addition, smudged blank slides had been added, at randomly determined positions, to each of the four blocks of stimuli. At the end of the experiment the subject was given a deck of index cards, on each of which was typed one of the neutral and one of the emotional words selected for him. The deck was randomly ordered and also included four pairs consisting of neutral words and four pairs consisting of emotional words. All words in all possible pairings were included in the deck. He was asked to "simply name one word from each card." From the initial meeting of experimenter and subject to a detailed, probing interview at the end of the experiment, all verbal behavior was tape recorded. Its evaluation by the experimenter was subsequently checked with the aid of a male and a female clinician.

As a result of the procedures which have been described, the selected emotion-arousing and neutral words were presented at a level at which the subjects frequently reported a "word-like" stimulus, sometimes confidently recognized stimuli, never discriminated blank slides from other difficult-to-discriminate stimuli, and obtained a frequency of accurate response significantly better than chance (usually about 30%). The determination of this level for presentation equated the subjects for general accuracy, which was not a variable of experimental interest. The practice it provided led to skilled and relatively stable performance (e.g., an absence of inappropriately timed blinks).

Probability Formulation of Results

The basis for the *perceptual-defense* measure was the difference between the probability of a neutral response which

was accurate and the probability of an emotional response which was accurate. As Goldiamond (1958) and others have demonstrated, this difference is conceptually equivalent to differences between ascending-method-of-limits thresholds since it is the "sufficiently high" probability of an accurate response which determines the point at which the ascent ends and the threshold is recorded. However, use of this difference, rather than a "threshold," aids in the comparison of perceptual defense with response bias.

The basis of the *response-bias* measures was the difference between the probability of a neutral response which was a guess and that of an emotional response which was a guess, determined within some group of responses which corresponded so little to a discriminative stimulus that the term guess seemed an accurate description, e.g., any response to a blank slide or the response "suck" to the stimulus "dance." When making such responses, the subjects sometimes commented "that was a pure guess."

The guesses used for estimating response bias included (a) responses to blanks, i.e., guesses in the absence of discriminative stimuli, (b) all inaccurate responses, i.e., guesses made in the presence of various unrecognized stimuli, (c) guesses naming stimuli never presented, i.e., particularly wild guesses, (d) guesses made when neutral words were presented, (e) guesses made when emotion-arousing words were presented, (f) choices made from clearly presented words. Three additional response bias measures were obtained by correcting *b*, *d*, and *e* for chance accuracy.²

² For each subject group, separate corrections were made for each stimulus-response combination (neutral responses to neutral stimuli, neutral responses to emotional stimuli, etc.). However the logic of

Thus the experiment yielded a group of response-bias measures obtained under a variety of stimulus conditions and based both on response groups about as large as the group consisting of accurate responses and on response groups nearly four times this large.

The basis of each measure was the difference between the probability of a neutral response and that of an emotional response within a given response subset, i.e., a difference between conditional probabilities. Results are conveniently written and communicated in terms of the differences between percentages.

Convergent validity for response-bias measures is indicated by correlations obtained in 52 male and female subjects

the procedure is the same in each case and is sufficiently illustrated by describing the correction of neutral responses to neutral stimuli. The formulas correct for the probability of chance accuracy, P (chance accuracy), which is adjusted for the subject's tendency to choose words from the presented group and for the subject's response bias. Corrected response frequencies are used in place of the observed frequencies in computing corrected measures of response bias.

1. Since only half the words on the subject's list are presented, he may increase his probability of chance accuracy by using these presented words in his guesses. The probability of chance accuracy is the joint probability of (a) accuracy when words in the presented group are used in guesses, and (b) choice from the presented group.

P (chance accuracy) = $1/n P$ (choice from presented group)

n = number of words in the presented group.

The probability of choice from the presented group is empirically determined from responses to blank slides.

2. The expected number of chance accuracies (A) when a neutral word is actually presented is

NP (chance accuracy) = A

N = number of guesses to neutral words.

3. However, the expected value of NP (chance accuracy) will also be increased (to A') by a propensity to choose neutral words

TABLE 1
INSIGNIFICANT RESPONSE BIAS

Response group for bias measure	Males	<i>t</i>	Females	<i>t</i>
Responses to blank slides	+3.44	.404	- 2.88	.574
All errors	+6.68	1.462	- 3.60	.651
Errors naming stimuli never presented	+7.28	1.000	- .14	.016
Responses to unrecognized neutral words	+6.92	.739	+ 1.16	.138
Responses to unrecognized emotional words	+5.78	1.067	-10.72	1.630
All errors, corrected	+6.38	1.837*	- 3.60	.724
Responses to unrecognized neutral words, corrected	+6.78	1.187	+ 1.30	.719
Responses to unrecognized emotional words, corrected	+6.06	1.338	- 8.48	1.601

* $p < .10$ using a two-tailed test, the only such result.

Note.—Bias equals percentage emotional response less percentage neutral response averaged for 26 males or females. Corrected measures were corrected for the effects of chance accuracy.

for which the amount of response bias was not statistically significant for either sex on any measure, i.e., there was little reason to expect a great amount of shared variance due to response bias. Bias demonstrated in responses to blanks and bias in responses to unrecognized words correlate (Pearson $r = .53$, $p < .001$). Bias in

(i.e., by response bias). A correction for chance accuracy, adjusted for response bias, is,

NP (chance accuracy) + c NP (chance accuracy) = A' .

$$c = P \text{ (neutral to blank)} - .50$$

Here, " c " is the proportion of neutral responses given when blanks are presented less .50—the proportion suggesting no response bias. Such an adjustment is appropriate whether or not the subject's bias is greater than .50. It will appropriately reduce his NP (chance accuracy) should he tend not to give neutral responses.

4. The corrected number of neutral guesses to neutral stimuli. NN (corrected), is the number of inaccurate neutral responses (NN) to neutral stimuli plus the number accurate by chance (A').

$$NN \text{ (corrected)} = NN + A'$$

In practice, correction did not make statistically significant changes in response-bias measures. Measures unaffected by chance accuracy indicate this was as it should have been.

responses to unrecognized stimuli and bias in choices from paired words also correlate significantly ($r = .26$, $p < .05$), although at the lower level expected because of the differences in stimulus conditions and instructions, which stress accuracy in the first case and response preference in the second. Bias in responses to blanks and that in choices have a correlation of .22 (.231 is significant at $p < .05$). Further information on the agreement of measures is given in Table 1, which shows the mean for each measure.

Results

Response Bias. Although stimuli, subjects, and tachistoscopic settings were chosen with the intent of obtaining perceptual defense, the response bias one might expect was absent. On all measures obtained during stimulus presentation, response bias was both slight and statistically insignificant (Table 1). Only bias in all errors, corrected for chance accuracy, approached being significantly different from zero. This is true only for percentage emotional response less percentage neutral response for males (+6.38%, $t = 1.837$, $p < .10$). One

TABLE 2
SIGNIFICANT PERCEPTUAL DEFENSE

	<i>M</i>	<i>t</i>
Males	-19.24	2.077**
Females	+17.34	2.157**
Sex	36.58	2.982***

** $p < .05$.

*** $p < .01$.

Note.—Perceptual defense is percentage emotional response less percentage neutral response averaged for the accurate responses of 26 males or females. Probabilities are two-tailed.

such approach to significance would probably be expected in any study yielding over 10 means. The difference between sexes is insignificant ($t = 1.547$). Free choices from paired, clearly presented stimuli also fail to demonstrate significant bias. Mean differences were $-.34\%$ for males, $+4.10\%$ for females. Neither value differs significantly from zero. The difference between sexes is statistically insignificant ($t = .765$).

Trends in the data suggest that males preferred emotional words. A female preference for neutral words is suggested by (a) the measures suggesting the most bias (responses to unrecognized emotional words), (b) the measure based on the most responses (all errors), and (c) most of the bias measures.

Perceptual Defense. According to a response-bias hypothesis, significant perceptual defense should be absent and any trends which are present should show males recognizing emotional words *more* easily and females recognizing them *less* easily than the neutral, control stimuli. However, significant perceptual defense is present for both males and females (who show a sex difference like that demonstrated by Brown).

For both males and females, the stimuli which are recognized most readily are the opposite of those pre-

dicted by a response-bias hypothesis. Of the four predictions possible, a response-bias hypothesis led to the one which is incorrect for both groups of subjects.

A response-bias hypothesis also suggests that bias in response to blanks (which do not restrict bias) will be equal to or greater than perceptual defense shown by responses to stimuli (which do restrict bias). Such an expectation follows directly from the notion that it is the characteristics of responses, rather than the emotion-arousing properties of stimuli, which cause perceptual defense; its test uses the response-bias most frequently mentioned in previous discussion and research (Goldiamond, 1958, 1960; Goldstein, 1962). However, because of the results of Mathews and Wertheimer (1958) an opposite prediction was made. Perceptual defense was expected to be "over-and-above" response bias. As predicted, perceptual defense proved significantly greater than response bias, both for males ($t = 1.734$, $p < .05$) and for females ($t = 2.010$, $p < .05$). In general, perceptual defense is greater than response bias on any measure.

Finally, the response-bias hypothesis implies a prediction of the direction of threshold differences. The suppression of socially unacceptable responses or the omission of unexpected responses could produce a false impression of "perceptual" defense predictable from response bias measures. However, in this study, response-bias and perceptual defense means generally differ in direction. For individual subjects, prediction of the direction of a recognition threshold difference from bias in response to blanks is wrong for 31 of the 52 subjects. Only 26 errors would be expected by chance.

None of the implications of a response-bias hypothesis are supported

by the study. In fact predictions made on the basis of pure chance would usually be successful more frequently.

TOWARD A MORE COMPLETE CONCEPTION OF ACCURATE RESPONSE TO EMOTION-AROUSING STIMULI

This article has criticized the oversimplification of definition and design found in early studies of perceptual defense and in more recent studies of a response-bias hypothesis.

When response bias and perceptual defense were measured nearly simultaneously, the implications of the response-bias hypothesis proved more wrong than right. Although subjects and stimuli were selected by objective criteria designed to increase the likelihood of perceptual defense (and it did prove present) response bias was absent. Both individual and group perceptual defense tended to be in a direction opposite to that which would be predicted by the hypothesis.

Further evaluation of other response-bias and perceptual defense studies will support these findings and suggest the need for a more complete conception of response tendencies in the perceptual defense experiment.

Response Bias

A serious, programmatic attack on the measurement of response tendencies has been provided by Goldstein and his associates (Goldstein, 1961, 1962; Goldstein, Himmelfarb, & Feder, 1962; Goldstein & Ratleff, 1961).

One study (Goldstein, 1961) indicates a relationship between associative reaction time and the subject's anxiety level, independent of the complexity of the habit hierarchy to the word. This finding, like the findings of Sarason (1959) and others, supports the use of a word association test as a technique for the selection of emotion-arousing stimuli. Of course, many artifacts

must be avoided (e.g., the contribution of the subject's mean reaction time, general familiarity of the word, familiarity within an appropriate norm group, the presence of words all subjects prefer not to discuss, etc.).

Using a word association test, Goldstein (1962) selected emotion-arousing words for which the subject's associative reaction time was longer than for structurally similar neutral words of equal frequency in English. Subjects were randomly assigned to (a) a group to which blank slides were repeatedly presented, (b) a group to which selected stimuli were very rapidly presented, or (c) a group to which stimuli were presented with regularly increasing clarity. All groups showed a bias against selected "emotional" stimuli. For the blank slide group, bias was not significantly less than for the group repeatedly presented the selected neutral and emotional words. In this comparison, male and female results were averaged.

Superficially inspected, the findings conflict with those reported by Blum (1955), Mathews and Wertheimer (1958), Minard (1963), and Nelson (1955). All report perceptual defense greater than response bias.

More carefully inspected, it is evident that Goldstein's study really serves as a useful critique of certain conceptions of perceptual defense.

Unpleasant and socially unacceptable stimuli were allowed among the selected emotion-arousing words and a bias against them is evident in the data. This significant response bias serves to invalidate the measure of perceptual defense.

This is probably especially true for females, who avoid saying the emotion-arousing words under all experimental conditions, although females recognized highly emotion-arousing words especially *readily* in studies by

Brown (1961) and Minard (1963). Even in Goldstein's data, mean female accuracy to emotional words is *greater* than mean male accuracy, although their "pseudo accuracy" (to blank slides) is significantly *less*. A pseudo accuracy was a correspondence of the subject's response with a word on the list used in the stimulus presentation session. The results for accurate responses of the male and female groups fail to predict the significant difference in pseudo accuracy to blank slides, since the sex differences go in opposite directions in the two groups of responses. Such a differential effect suggests that response bias *interacts* with stimulus characteristics to determine the probability of accurate responses when stimuli are actually presented.

Nevertheless, the procedures used follow directly from a popular (but restricted) definition of perceptual defense and from the implications of a response-bias hypothesis. Goldstein's study warrants comment not because its procedures are inferior to those typically used, but because they are unfortunately frequent and often not so well thought out or reported.

Averaging across subjects differing in sex and personality and across stimuli differing in emotion-arousing properties might be expected to average out *individual* differences in perception, leaving, as a residual, verbal response biases generally supported in the culture.

Altered Experimental Situations. The influence of response tendencies on perceptual defense has also been studied by changing the responses required of subjects. Results have been given perceptual-defense interpretations (e.g., Blum, 1954; Nelson, 1955) and have also been used to question them (Nothman, 1962; Postman, Bronson, & Gropper, 1953). Refutations of per-

ceptual defense have typically used socially unacceptable stimuli. The almost certain presence of response bias, when public report of such stimuli is required, obscures the meaning of results. For any stimuli, a change in instructions may well change the experimental situation, e.g., it may change the effective level of emotion aroused by the stimulus. Changes in the event to be reported may change the definition of recognition, e.g., when the subject states which stimulus was "more clear" or "stood out the most" his task may not involve recognition of the whole stimulus at all. It is interesting that on very brief presentation emotion-arousing stimuli are described as "more clear," whereas on longer presentation they are not (Mattson & Natsoulas, 1962). However, the early clarity of details of emotion-arousing stimuli might have an effect which either helps or hinders in the process of stimulus recognition. Understanding the effects of changes in the experimental situation may require the simultaneous measurement of perceptual defense, response bias, and "stimulus clarity." At present, changes caused by altering the experimental situation do not contribute greatly in understanding the contribution of response bias to perceptual defense, but they might complicate the task of this review. Therefore, when dealing with the relation of response bias to perceptual defense, this review uses perceptual defense studies (or parts of studies) in which subjects were instructed to achieve *accurate stimulus recognition* (i.e., to name the presented stimuli or to name a stimulus and its location if stimuli are presented simultaneously).

Perceptual Defense

Mathews and Wertheimer found perceptual defense greater than response bias in subjects high on the Hy

scale of the MMPI. Their study serves as an example of several studies similar in certain respects: All have (a) measured perceptual defense and response bias nearly simultaneously, (b) used responses naming presented stimuli as at least one measure of "perceptual defense contaminated with response bias," (c) used responses naming *unpresented* stimuli (expected by the subject) as at least one response-bias measure, (d) provided subjects with a small sample of possible responses from which they must try to make an accurate forced choice, and (e) used repeated stimulus presentations at a level allowing a better-than-chance probability of accurate response rather than an ascending method of limits terminated by *one* accurate response. These studies include those of Blum (1955), Mathews and Wertheimer (1958), Minard (1963), and Nelson (1955).

All have found evidence for perceptual defense, over and above the response bias which may contaminate a measure of perceptual defense. This independence of perception and indicator response may be present in other studies as well, but obscured by a failure to measure response bias or to control for its effects on perceptual defense measures.

The studies provide replications which check on possible artifacts and suggest generality for the findings. For example, Nelson (1955) replicated Blum (1955) but equated stimuli more carefully. Both simultaneously presented Blacky pictures in groups of four and asked subjects to name stimuli and their locations. Goldstein, Himmel-farb, and Feder (1962) have demonstrated "stimulus-location" bias (with words) but Nelson attempted to control for this with a preexperimental presentation of stimuli.

The author extended the results of

Mathews and Wertheimer (1958) to accurate responses by groups differing in sex. The following reanalysis of portions of his data provided more direct replication and support. The difference between percentage of neutral response and percentage of emotional response was obtained for responses naming words in the group presented (excluding a group of grossly inaccurate stimulus response combinations so defined before the experiment). A reliability coefficient (correlating half these responses with the remainder of them) was computed for the 52 males and females of this study. It was .38 ($p < .005$), suggesting a moderate but significant reliability despite the presence of both response bias and stimulus recognition. Using a nonparametric Wilcoxon T (Siegel, 1956) a significant difference was found between this measure and bias in responses to blank slides for both males (22.68%, $p < .01$) and for females (20.32%, $p < .05$). Within the sample of responses to blanks, the difference between bias in responses naming presented words and bias in responses naming unpresented words was not significant for males ($t = .156$) or females ($t = .381$). Because of the nature of the measures used in both the Mathews and Wertheimer study and its replication, the "contamination" by response bias could not in itself raise the perceptual defense measure above the level of response bias measured independently. It was the characteristics of presented stimuli, not of responses, which produced perceptual defense over and above subjects' response bias.

Conclusions

An overall interpretation may be made on the basis of response-bias studies such as that by Goldstein (1962) and perceptual defense studies such as that reported here. The prob-

ability of an accurate response to a stimulus may be a result of the interaction of (a) bias against certain responses and (b) emotion-arousing properties of the presented stimuli. By appropriate experimental design, the effect of either variable may be greatly reduced. It is an interesting possibility that, due to differences in operant response conditioning or to conditioned emotional behavior, the effects of these variables may differ from person to person. Most important, in the light of current controversy, is the evidence that perceptual defense does occur, independent of the nature or presence of response bias.

REFERENCES

- ALLPORT, F. H. *Theories of perception and the concept of structure*. New York: Wiley, 1955.
- Blackwell, H. R. Psychophysical thresholds: Experimental studies of methods of measurement. Ann Arbor: University of Michigan Press, *Engineering Research Bulletin* No. 36, cited by W. N. Dember, *Psychology of Perception*, New York: Holt, Rinehart, & Winston, 1960. Pp. 43-44, 114-115.
- BLUM, G. S. An experimental reunion of psychoanalytic theory with perceptual vigilance and defense. *Journal of Abnormal and Social Psychology*, 1954, 49, 94-98.
- BLUM, G. S. Perceptual defense revisited. *Journal of Abnormal and Social Psychology*, 1955, 51, 24-29.
- BROWN, W. P. Conceptions of perceptual defense. *British Journal of Psychology*, *Monogr. Supp.*, 1961, No. 35.
- BRUNER, J. S., & POSTMAN, L. Emotional selectivity in perception and reaction. *Journal of Personality*, 1947, 66, 69-77.
- BRUNER, J. S., & POSTMAN, L. Symbolic value as an organizing factor in perception. *Journal of Social Psychology*, 1948, 27, 203-208.
- CHODORKOFF, B., & CHODORKOFF, J. Perceptual defense: An integration with other research findings. *Journal of General Psychology*, 1958, 58, 75-80.
- CLAPP, C. D. Two levels of unconscious awareness. Unpublished doctoral dissertation, University of Michigan, 1951.
- DAHLSTROM, W. G., & WELSH, G. S. *An MMPI handbook*. Minneapolis: Univer. Minnesota Press, 1960.
- DEESE, J., LAZARUS, R. S., & KEENAN, J. Anxiety, anxiety reduction and stress in learning. *Journal of Experimental Psychology*, 1953, 46, 55-60.
- DELUCIA, J. J., & STAGNER, R. Emotional vs. frequency factors in word-recognition time and association time. *Journal of Personality*, 1953, 22, 299-309.
- DULANY, D. E. Avoidance learning of perceptual defense and vigilance. *Journal of Abnormal and Social Psychology*, 1957, 55, 333-338.
- ERIKSEN, C. W. Perceptual defense as a function of unacceptable needs. *Journal of Abnormal and Social Psychology*, 1951, 46, 557-564. (a)
- ERIKSEN, C. W. Some implications for TAT interpretation arising from need and perception experiments. *Journal of Personality*, 1951, 19, 283-289. (b)
- ERIKSEN, C. W. Defense against ego-threat in memory and perception. *Journal of Abnormal and Social Psychology*, 1952, 47, 230-235. (a)
- ERIKSEN, C. W. Individual differences in defensive forgetting. *Journal of Experimental Psychology*, 1952, 44, 442-446. (b)
- ERIKSEN, C. W. The case for perceptual defense. *Psychological Review*, 1954, 61, 175-182. (a)
- ERIKSEN, C. W. Psychological defenses and "ego strength" in the recall of completed and incompleting tasks. *Journal of Abnormal and Social Psychology*, 1954, 49, 45-50. (b)
- ERIKSEN, C. W. Some personality correlates of stimulus generalization under stress. *Journal of Abnormal and Social Psychology*, 1954, 49, 562-566. (c)
- ERIKSEN, C. W. Discrimination and learning without awareness: A methodological survey and evaluation. *Psychological Review*, 1960, 67, 279-300.
- ERIKSEN, C. W., & DAVIDS, A. The meaning and clinical validity of the Taylor manifest anxiety scale and the Hysteria-Psychasthenia scales of the MMPI. *Journal of Abnormal and Social Psychology*, 1955, 50, 135-137.
- ERIKSEN, C. W., & LAZARUS, R. S. Perceptual defense and projective tests. *Journal of Abnormal and Social Psychology*, 1952, 47, 302-308.
- EYSENCK, H. J. The concept of statistical significance and the controversy about one-

- tailed tests. *Psychological Review*, 1961, 67, 269-271.
- EYSENCK, H. J., & CLARIDGE, G. The position of hysterics and dysthymics in a two-dimensional framework of personality description. *Journal of Abnormal and Social Psychology*, 1962, 64, 46-55.
- FENICHEL, O. *The psychoanalytic theory of neurosis*. New York: Norton, 1945.
- FRANKMAN, JUDITH P., & ADAMS, J. A. Theories of vigilance. *Psychological Bulletin*, 1962, 59, 257-272.
- FREEMAN, J. T. Set or perceptual defense? *Journal of Experimental Psychology*, 1954, 48, 283-288.
- FREEMAN, J. T. Set versus perceptual defense: A confirmation. *Journal of Abnormal and Social Psychology*, 1955, 51, 710-712.
- FREUD, ANNA. *The ego and the mechanisms of defence*. London: Hogarth Press, 1937.
- GOLDIAMOND, I. Indicators of perception: I. Subliminal perception, subception, unconscious perception: An analysis in terms of psychophysical indicator methodology. *Psychological Bulletin*, 1958, 55, 373-411.
- GOLDIAMOND, I. Word frequency, accuracy of recognition, and conditioning: or just what role does a discriminative stimulus play in a recognition experiment. Paper presented at symposium: Word frequency as a variable in behavioral studies. American Psychological Association, Chicago, September 1960.
- GOLDIAMOND, I., & HAWKINS, W. F. Vexiersuch: The log relationship between word-frequency and recognition obtained in the absence of stimulus words. *Journal of Experimental Psychology*, 1958, 56, 457-463.
- GOLDSTEIN, M. J. A relationship between anxiety and oral word association performance. *Journal of Abnormal and Social Psychology*, 1961, 62, 468-471.
- GOLDSTEIN, M. J. A test of the response probability theory of perceptual defense. *Journal of Experimental Psychology*, 1962, 63, 23-28.
- GOLDSTEIN, M. J., HIMMELFARB, S., & FEDER, WALDA. A further study of the relationship between response bias and perceptual defense. *Journal of Abnormal and Social Psychology*, 1962, 64, 56-62.
- GOLDSTEIN, M. J., & RATLEFF, J. Relationship between frequency of usage and ease of recognition with response bias controlled. *Perceptual and Motor Skills*, 1961, 13, 171-177.
- HICK, W. E. A note on one-tailed and two-tailed tests. *Psychological Review*, 1952, 59, 316-318.
- HOWES, D., & SOLOMON, R. L. A note on McGinnies' "Emotionality and perceptual defense." *Psychological Review*, 1950, 57, 229-234.
- HOWES, D., & SOLOMON, R. L. Visual duration threshold as a function of word probability. *Journal of Experimental Psychology*, 1951, 41, 401-410.
- HOWIE, D. Perceptual defense. *Psychological Review*, 1952, 59, 308-315.
- INGLIS, J. Abnormalities of motivation and "ego-functions." In H. J. Eysenck (Ed.), *Handbook of abnormal psychology*. New York: Basic Books, 1961.
- KRASNER, L. Studies of the conditioning of verbal behavior. *Psychological Bulletin*, 1958, 55, 148-170.
- LACEY, O. W., LEWINGER, N., & ADAMSON, J. F. Foreknowledge as a factor affecting perceptual defense and alertness. *Journal of Experimental Psychology*, 1953, 45, 169-174.
- LAZARUS, R. S. Is there a mechanism of perceptual defense? A reply to Postman, Bronson, and Gropper. *Journal of Abnormal and Social Psychology*, 1954, 49, 306-398.
- LAZARUS, R. S., & LONGO, N. The consistency of psychological defenses against threat. *Journal of Abnormal and Social Psychology*, 1953, 48, 495-499.
- MCGINNIES, E. Emotionality and perceptual defense. *Psychological Review*, 1949, 56, 244-251.
- MATHEWS, ANN, & WERTHEIMER, M. A "pure" measure of perceptual defense uncontaminated by response suppression. *Journal of Abnormal and Social Psychology*, 1958, 57, 373-376.
- MATTSON, J. M., & NATSOULAS, T. Emotional arousal and stimulus duration as determinants of stimulus selection. *Journal of Abnormal and Social Psychology*, 1962, 65, 142-144.
- MILLER, CRISTINE. Consistency of cognitive behavior as a function of personality characteristics. *Journal of Personality*, 1954, 23, 233-249.
- MINARD, J. G. The measurement and conditioning of "perceptual defense" and response suppression. Unpublished doctoral dissertation, University of Colorado, 1963.
- NELSON, S. E. Psychosexual conflicts and defenses in visual perception. *Journal of Abnormal and Social Psychology*, 1955, 51, 427-433.
- NOTHMAN, F. H. The influence of response conditions on recognition thresholds for tabu words. *Journal of Abnormal and Social Psychology*, 1962, 65, 154-161.

- POSTMAN, L., BRONSON, WANDA C., & GROPPER, G. L. Is there a mechanism of perceptual defense? *Journal of Abnormal and Social Psychology*, 1953, 48, 215-224.
- SALZINGER, K. Experimental manipulation of verbal behavior: A review. *Journal of General Psychology*, 1959, 61, 65-94.
- SARASON, I. G. Relationship of measures of anxiety and experimental instructions to word association performance. *Journal of Abnormal and Social Psychology*, 1959, 59, 37-42.
- SARASON, I. G., & MINARD, J. G. Test anxiety, experimental instructions, and the Wechsler Adult Intelligence Scale. *Journal of Educational Psychology*, 1962, 53, 299-302.
- SARASON, I. G., & MINARD, J. G. Interrelationships among subjects, experimenter, and situational variables. *Journal of Abnormal and Social Psychology*, 1963, 67, 87-92.
- SIEGEL, S. *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill, 1956.
- SPENCE, D. P. A new look at vigilance and defense. *Journal of Abnormal and Social Psychology*, 1957, 54, 103-108.
- SPENCE, JANET T. Contribution of response bias to recognition thresholds. *Journal of Abnormal and Social Psychology*, 1963, 66, 339-345.
- TAYLOR, J. A., ROSENFELDT, D. C., & SCHULTZ, R. W. The relationship between word frequency and perceptibility with a forced choice technique. *Journal of Abnormal and Social Psychology*, 1961, 62, 491-496.
- TRUAX, C. B. The repression response to implied failure as a function of the Hysteria-Psychasthenia index. *Journal of Abnormal and Social Psychology*, 1957, 55, 574-578.
- WALL, H. W., & GUTHRIE, G. M. Academic stress and perceptual threshold. *Journal of General Psychology*, 1959, 61, 269-273.
- ZAJONC, R. B. Response suppression in perceptual defense. *Journal of Experimental Psychology*, 1962, 64, 206-214.

(Received December 20, 1963)

PSYCHOLOGICAL REVIEW

PRIMARY MEMORY¹

NANCY C. WAUGH

Harvard Medical School

AND DONALD A. NORMAN

Center for Cognitive Studies, Harvard University

A model for short-term memory is described and evaluated. A variety of experimental data are shown to be consistent with the following statements. (a) Unrehearsed verbal stimuli tend to be quickly forgotten because they are interfered with by later items in a series and not because their traces decay in time. (b) Rehearsal may transfer an item from a very limited primary memory store to a larger and more stable secondary store. (c) A recently perceived item may be retained in both stores at the same time. The properties of these 2 independent memory systems can be separated by experimental and analytical methods.

It is a well-established fact that the longest series of unrelated digits, letters, or words that a person can recall verbatim after one presentation seldom exceeds 10 items. It is also true, however, that one can nearly always recall the most recent item in a series, no matter how long the series—but only if this item may be recalled immediately, or if it may be rehearsed during the interval between its presentation and recall. Otherwise it is very likely to be lost. If we may assume that attending to a current item precludes reviewing a prior one, we can say that

the span of immediate memory must be limited in large part by our inability to rehearse, and hence retain, the early items in a sequence while attempting to store the later ones. Our limited memory span would then be but one manifestation of our general inability to think about two things at the same time.

Why should an unrehearsed item in a list be forgotten so swiftly? Is its physiological trace in some sense written over by the traces of the items that follow it? Or does this trace simply decay within a brief interval, regardless of how that interval is filled? Tradition, in the guise of interference theory, favors the first explanation (McGeoch, 1932; Postman, 1961), although some psychologists now think that new memory traces must fade autonomously in time (Brown, 1958; Conrad, 1957; Hebb, 1949). Until now, no one has reported any data

¹ This research was supported in part by Research Grants No. MH 05120-02 and MH 08119-01 from the National Institutes of Health, United States Public Health Service, to Harvard University, Center for Cognitive Studies and to Harvard Medical School, respectively. The second author was supported by a National Science Foundation Postdoctoral Fellowship at the Center for Cognitive Studies.

which clearly contradict either of these ideas. In fact, when we first considered the problem of the instability of recent memory traces, we thought it entirely possible that both decay and interference operate over brief retention intervals to produce forgetting, and we therefore designed an experiment to weigh their respective effects. The results of this experiment were unexpectedly straightforward—and seemingly inconsistent with certain other existing data on immediate retention. We have been able, however, to formulate a simple quantitative model which relates our results to those reported by other investigators. What began as an attempt to evaluate two very general hypotheses about the forgetting of recent events has therefore resulted in a specific theory of short-term memory.

We shall describe our experiment in Section I below. A major portion of this paper, Section II, will be concerned with the description and application of our model. In Section III we shall discuss this model in relation to the general question of whether short- and long-term retention represent distinguishably different psychological processes.

I. PROBE-DIGIT EXPERIMENT

Our experiment was designed to measure the recall of a minimally rehearsed verbal item as a joint function of the number of seconds and the number of other items following its presentation. The general procedure was as follows. Lists of 16 single digits were prepared with the aid of a standard table of random numbers, under the constraint that no digit should appear more than twice in a row. The last digit in every list was one that had occurred exactly once before, in Position 3, 5, 7, 9, 10, 11, 12, 13, or 14. On its second appearance, this "probe-digit" was a cue for the recall of the digit that had followed it initially.

The lists were recorded on two magnetic tapes; they were read in a monotone voice by a male speaker at a constant rate of

either one or four digits per second. Each of the nine possible probe-digit positions was tested 10 times. The two tapes accordingly contained 90 test lists (plus 8 practice lists) apiece, all read at the same rate. The last digit in every list, the probe-digit, was accompanied by a high-frequency tone to aid the subject in detecting the end of the list. The position of the initial presentation of the probe varied randomly from list to list on each of the two tapes.

The subject's task was to write down the digit that had followed the probe digit in the list, guessing if he did not know. Since the probe-digit was unique in Positions 1 through 15, there was only one possible correct answer on any trial. Every subject listened to the list through earphones for a total of 12 experimental sessions, 6 with each tape, alternating between fast and slow lists. The first session under each condition and the first eight lists listened to in each session were considered to be practice and, unknown to the subject, were not scored.

The subjects received explicit instructions to control rehearsal by "thinking only of the last digit you have heard and never of any of the earlier ones." These instructions were repeated before the second session, and occasional reminders were given throughout the course of the experiment. Thus, the subjects were to rehearse every item during the interitem interval immediately following it. Our instructions were not designed to eliminate the rehearsal of single items as such, but rather to eliminate the rehearsal of *groups* of digits. The experiment actually tested the retention of a digit pair, the probe-digit and its successor. The retention of this pair should be independent of the interitem interval, if the instructions to avoid grouping were followed faithfully. We hoped, in effect, to test the retention of unrehearsed pairs of digits under two rates of presentation.

The subjects were four Harvard undergraduates, three males and one female.

The responses were scored and analyzed to yield a serial position curve for each rate of presentation, relating the relative frequency of an item's correct recall to its distance from the end of the list. A comparison of the two functions allows us to assess the relative effects of decay and interference on short-term forgetting, according to the following line of reasoning. Consider the recall of Item i from the end of the line. If the list was read at the rate of one item per second, then i items would have intervened, and i seconds would have elapsed between

the time the subject heard the item and the time he attempted to recall it. (We count the second appearance of the probe-digit as an intervening event.) If the items were read at the rate of four per second, on the other hand, then only $i/4$, rather than i , seconds would have elapsed between the occurrence of Item i and the subject's attempt to recall it. A total of i other items would, of course, still have intervened between these two events. Therefore, if the probability of recalling Item i from the end of a slow list were identical with the probability of recalling Item $4i$ from a fast list, we could conclude that recent memory traces decay in time, independently of one another. Conversely, if the probability of recalling Item i were invariant with rate of presentation, we could conclude that rapid forgetting is caused primarily by retroactive interference.

The results of the experiment are shown in Figure 1. The scores for the individual subjects are presented in Figures 1A and 1B. The pooled data, corrected for guessing, are shown in Figure 1C.² Each point in Figures 1A and 1B is based on 50 observations; each point in Figure 1C, on 200. It is evident that there are consistent differences among subjects, but little interaction between subjects and serial positions. Furthermore, although there appears to be a slight interaction between relative frequency of recall, or $R(i)$, and rate of presentation, it is clear that the effect of rate is relatively small compared to the effect of serial position. The main source of forgetting in our experiment was interference.

The differences between the two sets of points shown in Figure 1C are not statistically reliable, according to an analysis of variance performed on the number of items recalled by each sub-

²The response set—the 10 digits—was known to the subjects, and they knew that the probe would not be the same as the test digit. Thus the probability of correctly guessing the answer, g , was $1/9$. A standard normalizing technique was used to eliminate the effects of guessing from the data, namely, $p(\text{recall}) = [p(\text{correct}) - g]/(1 - g)$.

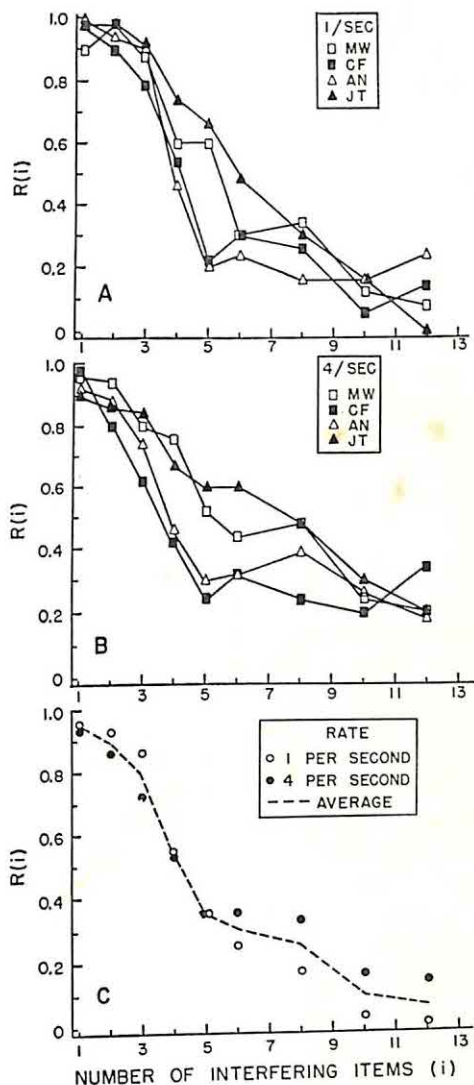


FIG. 1. Results of the probe-digit experiment. (Figures 1A and 1B represent retention functions for individual subjects under two rates of presentation; in Figure 1C these data have been pooled.)

ject at each value of i under the two rates of presentation ($F < 1$ for the mean square between rates tested against the interaction between subjects and rates). This conclusion is borne out by the results of nine Kolmogorov-Smirnov two-sample tests, one for each value of i , performed on the distributions of number of items recalled per subject per session under

the two rates of presentation. We have therefore fitted the points shown in Figure 1C with a function that represents the probability of recalling Item i from the end of a series, estimated across rates of presentation. This function decreases monotonically with i , attaining a value of about .07 at $i = 12$.

II. MODEL FOR PRIMARY MEMORY

When we compared the foregoing results with the typical outcome of the first trial in a standard list-learning experiment, we found ourselves facing two dilemmas. In the first place, it often happens that an item in a long list is recalled after 10 or 20, or even more, items have followed it. But in our experiment, probability of recall was effectively zero for the eleventh item in from the end of a list. In the second place, various investigators have shown that probability of recall increases with presentation time (see Posner, 1963), yet in our experiment this probability, for all practical purposes, was independent of the rate at which the digits were read.

In seeking for a way to account for these discrepancies, it occurred to us that one difference between our experiment and previous ones in this area is that we instructed our subjects not to think about any item in a list once the next had been presented. This instruction to avoid rehearsal is, to be sure, rather unorthodox, although not completely without precedent (Underwood & Keppel, 1962). In order to minimize rehearsal, many experimenters try to keep the subject so busy that he does not have time to rehearse; but we think it highly likely that a well-motivated subject who is trying to learn a list will rehearse unless specifically enjoined from doing so. The typical subject's account of how he learns a list (Bugelski, 1962; Clark, Lansford, &

Dallenbach, 1960) bears us out on this point. In fact, it is probably very difficult *not* to rehearse material that one is trying to memorize.

We shall assume here that rehearsal simply denotes the recall of a verbal item—either immediate or delayed, silent or overt, deliberate or involuntary. The initial perception of a stimulus probably must also qualify as a rehearsal. Obviously a very conspicuous item or one that relates easily to what we have already learned can be retained with a minimum of conscious effort. We assume that relatively homogeneous or unfamiliar material must, on the other hand, be deliberately rehearsed if it is to be retained. Actually, we shall not be concerned here with the exact role of rehearsal in the memorization process. We are simply noting that, in the usual verbal-learning experiment, the likelihood that an item in a homogeneous list will be recalled tends to increase with the amount of time available for its rehearsal. The probe-digit experiment has shown, conversely, that material which is not rehearsed is rapidly lost, regardless of the rate at which it is presented. It is as though rehearsal transferred a recently perceived verbal item from one memory store of very limited capacity to another more commodious store from which it can be retrieved at a much later time.

We shall follow James (1890) in using the terms *primary* and *secondary memory* (PM and SM) to denote the two stores. James defined these terms introspectively: an event in PM has never left consciousness and is part of the psychological present, while an event recalled from SM has been absent from consciousness and belongs to the psychological past. PM is a faithful record of events just perceived; SM is full of gaps and distortions. James believed that PM extends over a fixed

period of time. We propose instead that it encompasses a certain number of events regardless of the time they take to occur. Our goal is to distinguish operationally between PM and SM on the basis of the model that we shall now describe.

Consider the general scheme illustrated in Figure 2. Every verbal item that is attended to enters PM. As we have seen, the capacity of this system is sharply limited. New items displace old ones; displaced items are permanently lost. When an item is rehearsed, however, it remains in PM, and it may enter into SM. We should like to assume, for the sake of simplicity, that the probability of its entering SM is independent of its position in a series and of the time at which it is rehearsed. Thus, it would not matter whether the item was rehearsed immediately on entering PM or several seconds later: as long as it was in PM, it would make the transition into SM with fixed probability. (Our PM is similar to Broadbent's, 1958, *P* system. One difference between our two systems is that ours relates rehearsal to longer term storage, whereas his does not.)

Finally, we shall assume that response-produced interference has the same effect on an item in PM as does stimulus-produced interference. That is, the probability that an item in PM will be recalled depends upon (a) how many new items have been perceived plus (b) how many old ones have been recalled between its presentation and attempted recall. Thus, if an item appears in Position n from the end of a list and the subject attempts to recall it after recalling m other items, it is as if the item had appeared in position $i = n + m$ in the list, and recall was attempted at the end of the list. This assumption is rather strong, but recent studies by

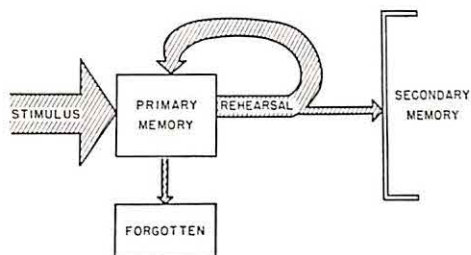


FIG. 2. The primary and secondary memory system. (All verbal items enter PM, where they are either rehearsed or forgotten. Rehearsed items may enter SM.)

Murdock (1963) and by Tulving and Arbuckle (1963) have, in fact, failed to reveal any consistent differences between stimulus- and response-induced interference in the retention of paired associates. It may not be unreasonable to suppose, therefore, that the two sources of interference exert equivalent effects on free and serial recall.

According to our hypothesis, then, the probability of recalling an item which has been followed by i subsequent items is given by the probability that it is in PM, in SM, or in both. Assuming that these probabilities combine independently,

$$R(i) = P(i) + S(i) - P(i)S(i) \quad [1]$$

where $R(i)$ is the probability that Item i will be recalled, $P(i)$ is the probability that it is in PM, and $S(i)$ the probability that it is in SM. The probability that this item is in PM is then given by

$$P(i) = [R(i) - S(i)] / [1 - S(i)]. \quad [2]$$

We assume that $P(i)$ is a monotonic decreasing function of i and that

$$\lim_{i \rightarrow \infty} P(i) = 0.$$

We should like specifically to test the hypothesis that $P(i)$ is independent of the value of $S(i)$ and, in fact, varies with i in the manner of the probe-digit data. (This hypothesis is stated more

formally in the Appendix.) In order to do so, we need data on verbal retention that meet the following requirements.

1. They should come from an experimental situation where at least some of the items are retrieved from PM.

2. The subject should have been allowed to rehearse, so that $S(i) > 0$.

3. The value of $S(i)$ should preferably be constant and independent of i .

4. The experimental lists should be long enough to let us estimate $S(i)$ for $i > 12$.

5. We should know the location of a given item in the stimulus list (n) and in the recall list (m), so as to be able to estimate the total number of interfering items ($i = n + m$).

Free Recall

The free-recall experiment is well suited to our purposes. Subjects can (and usually do) recall the last few items in a list right away, and the middle portion of the serial position curve (after the first three and before the last seven items) is effectively flat, thereby providing a convenient estimate of $S(i)$ (Deese & Kaufman, 1957; Murdock, 1962; Waugh, 1962).

Testing our hypothesis against data collected in a free-recall experiment therefore involves the following steps:

1. First, we estimate $S(i)$ from the average proportion of items recalled from the middle of a long list.

2. We then estimate $P(i)$ for each of the last seven items in the list by Equation 2.

3. We plot this estimate against $n + m = i$ and compare the resulting function with that shown in Figure 1.

Fortunately, we did not have to perform a free-recall experiment especially for this purpose: several such studies

have been carried out and reported in sufficient detail to enable us to test our hypothesis against their results. We have chosen to analyze four sets of data collected by three different investigators: Deese and Kaufman (1957), Murdock (1962), and two as yet unpublished experiments conducted by Waugh. The two principal variables that affect $S(i)$ in free recall appear to be length of list (the amount of material that is to be retained) and presentation time (the amount of time available for the rehearsal of a given item). Manipulating these variables results in orderly changes in the value of $S(i)$, so that our estimates range from .08 to .45 across the four experiments.

1. In Deese and Kaufman's study, the subjects listened to lists of 32 unrelated English words read at a rate of one per second, and began recalling them immediately after the last had been spoken. Deese and Kaufman have presented a serial position curve based on these data and have also reported the relation between an item's serial position in recall and its position in the original list. We can thereby estimate i for each item in their lists, letting an item's average position in recall be our estimator of the amount of response interference (m).³ We estimated $S(i)$ by the proportion of items recalled after the first three and before the last seven serial positions in the original list.⁴ (This

³ It is not really correct to use the average of the serial positions in recall as an estimate of $m + 1$: the total effect of response interference should depend on the variance of this distribution as well as on its mean or median. It is the only alternative open to us, however, since our correction for asymptote must be applied to the average proportion of items retained, estimated across serial position in recall.

⁴ In estimating $S(i)$, we ignored the recall of the first three items on a list because

same general procedure will be followed in our subsequent analyses.)

The last seven points of Deese and Kaufman's serial position curve, taken from their Figure 1 and corrected for asymptote according to Equation 2, are plotted as a function of i in Figure 3. The dashed lines in Figure 3 represent the 99% confidence limits for the probe-digit function: a standard error for each point was estimated across subjects and experimental sessions. The uncorrected data are shown in Table 1.

2. Waugh's experiments were concerned with determining the number of items freely recalled from long lists as a function of presentation time. In her first experiment, 24, 30, 40, 60, or 120 different monosyllabic English words were read to the subjects at a rate of one per second. The proportion of items recalled varied inversely with list length, so that for each length of list there is a different serial position function. The asymptotes of these functions range from approximately .08 to .20. Median serial position in recall ($m + 1$) was calculated for each of the last six items in a list;

they invariably show a primacy effect, perhaps the result of selective attention and rehearsal.

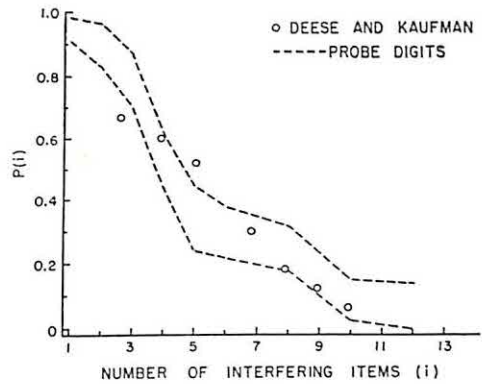


FIG. 3. Free-recall data from Deese and Kaufman (1957), corrected for asymptote and response interference.

Figure 4 shows $S(i)$ as a function of i for each of these items. The uncorrected data appear in Table 2.

In Waugh's second experiment, the subjects listened to 30 different words presented at a rate of 1, 2, 3, 4, or 6 seconds per word. In each case the presentations were either massed—that is, each word was read one, two, or three times in a row, at a rate of one word per second or of one word every two seconds—or they were distributed—each word was read once at one, two, three, four, or six different places in a list, at a rate of one word per second. The results of this experiment indicate that whether the repetitions are massed

TABLE 1
PROPORTION OF ITEMS FREELY RECALLED AS A FUNCTION
OF SERIAL POSITION AND TOTAL TIME PER LIST

Number of intervening items	List length \times seconds per item						
	32 \times 1	40 \times 1	20 \times 2	30 \times 1	15 \times 2	20 \times 1	10 \times 2
0	.72	.96	.95	.97	.97	.96	.95
1	.67	.85	.88	.89	.88	.84	.83
2	.60	.71	.75	.74	.80	.76	.71
3	.42	.51	.57	.52	.62	.62	.67
4	.32	.40	.43	.39	.58	.39	.58
5	.27	.27	.38	.33	.49	.30	.45
6	.22	.22	.38	.24	.42	.26	.45
6 + ^a	.17	.12	.27	.19	.38	.15	.45

Note.—Deese and Kaufman (1957), Column 1; Murdock (1961), Columns 2-6.

^a Entries in this row represent the asymptotic value of $R(n)$.

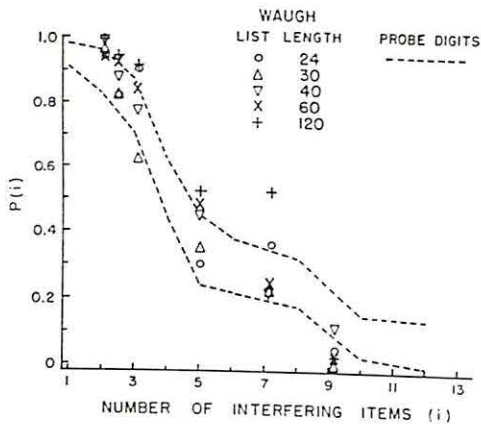


FIG. 4. Free-recall data from Waugh corrected for asymptote and response interference.

or distributed is of no importance; the probability that a word will be recalled is determined simply by the total number of seconds for which it is presented. Since this probability increases as a negatively accelerated function of presentation time, the asymptotic values of the serial position function obtained in this experiment ranged from approximately .14 (for 30 words each read once) to .45 (for 30 words each read six times). Average serial position in recall was again calculated for each of the last six items in a list. The retention functions for massed and distributed repetitions, cor-

rected for asymptote and response interference, are shown in Figures 5 and 6, respectively, along with the PM function obtained in our probe-digit experiment. The uncorrected data are shown in Table 2.

3. In Murdock's experiment, the subjects listened to lists of 20, 30, or 40 words read at a rate of 1 word per second and to lists of 10, 15, and 20 words read at a rate of 1 word every 2 seconds. Murdock found, as has Waugh (1963), that the probability of recalling a word that has been

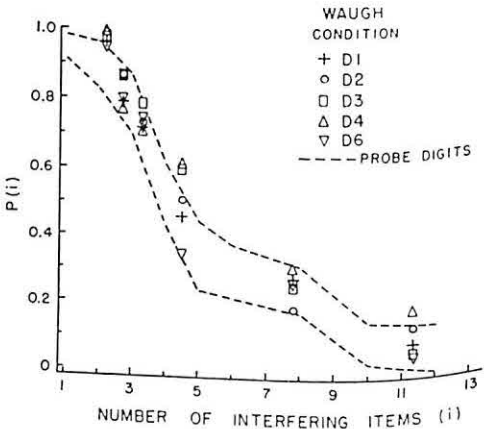


FIG. 5. Free-recall data from Waugh corrected for asymptote and response interference (1-6 distributed presentations per word).

TABLE 2
PROPORTION OF ITEMS FREELY RECALLED AS A FUNCTION OF STIMULUS INTERFERENCE AND NUMBER OF ITEMS PER LIST

Number of intervening items	List length				
	24	30	40	60	120
0	.95	.97	1.00	.95	1.00
1	.85	.85	.90	.93	.95
2	.92	.69	.81	.86	.92
3	.42	.46	.51	.53	.57
4	.47	.35	.31	.32	.57
5	.21	.17	.22	.14	.14
5 + ^a	.15	.17	.16	.12	.08

^a Entries in this row represent the asymptotic value of $R(n)$.

listened to for 2 seconds is almost exactly twice the probability of recalling a word that has been listened to for 1 second. Murdock's data can therefore be grouped into three pairs of serial position curves: 10 words read at a rate of 1 every 2 seconds versus 20 words read at a rate of 1 per second; 15 words read at a rate of 1 every 2 seconds versus 30 read at a rate of 1 per second; and 20 words read at a rate of 1 every 2 seconds versus 40 read at a rate of 1 per second. Within each pair, there are two asymptotes, one of which is approximately twice the value of the other.

We have corrected Murdock's curves for asymptote—that is, for $S(i)$ —and since he did not calculate serial position in recall for his words, we have plotted these corrected values of $P(i)$ against the average values of i calculated by Waugh for words recalled under similar conditions in the experiment just described (see Figures 5 and 6).⁵ Murdock's uncorrected data are shown in Table 1.

It is clear that an appreciable number of the points displayed in Figures 3 through 7 fall outside the confidence limits we have set for the probe-digit function. In general, the discrepancies between theoretical and observed values of $P(i)$ appear to be unsystematic. They may have resulted from either of two possible sources which would not be reflected in the variance of the probe-digit function.

In the first place, we assume that $S(i)$ is constant for all i . While $S(i)$ does not in fact seem to vary systematically with i in the middle of a list, individual words do differ greatly in their susceptibility of storage in secondary memory: the serial position function for free recall is haphazardly jagged rather than perfectly flat. Thus, even one anomalously easy word in Location n , for instance, can greatly inflate our estimate of $R(n)$ and hence $P(n)$. The probe-digit data would presumably not be subject to this kind of variability.

A second source of errors may lie in our estimation of i , or $m + n$. We have used average position in recall—call it $\bar{m} + 1$ —as our estimate of $m + 1$. Even a small error in this estimate can lead to a sizable discrepancy between a theoretical and an observed value of $P(i)$, especially around the steep early portion of the function.

⁵ The asymptotes for Murdock's curves were obtained by complementing his tabulated values for v (shown in his Table 2).

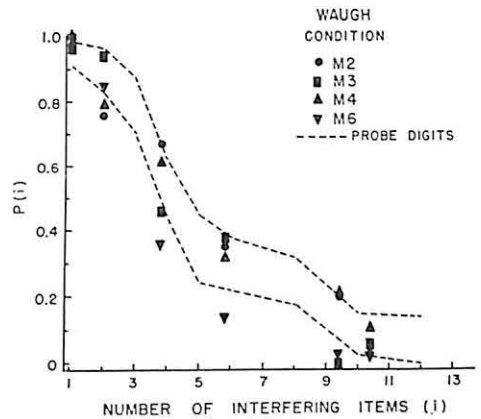


FIG. 6. Free-recall data from Waugh corrected for asymptote and response interference (2-6 massed presentations per word).

Errors of this sort would be reflected in Figures 4-7, where i and $P(i)$ are derived from either partially or completely independent sets of data (in Figures 4-6 and Figure 7, respectively). Furthermore, we should in any case expect some discrepancy on purely mathematical grounds between $P(i)$, where (i) is the mean of a point distribution, as in the probe-digit experiment, and $P(\bar{m} + n)$, where m can assume any of a number of values, as in the free-recall data we have analyzed. Unfortunately, we are un-

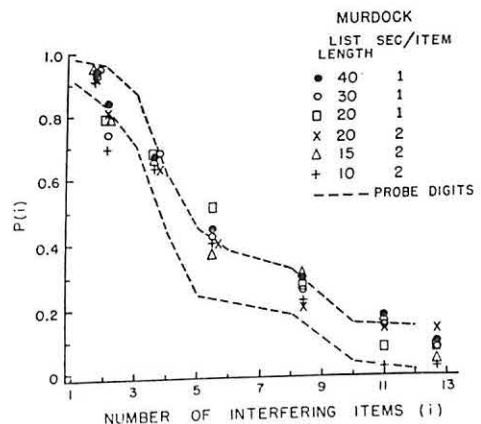


FIG. 7. Free-recall data from Murdock (1961), corrected for asymptote and response interference.

able to specify the magnitude of this expected discrepancy.

In view, therefore, of the likelihood of the errors we have just described, we believe that the fit between the probe-digit function and the free-recall data is fairly good and is, in fact, probably too close to attribute to chance. Actually, in one respect it is surprising that the probe-digit function should describe the free-recall data as well as it does. The probe-digit experiment tested the retention of digit pairs, whereas the free-recall experiments tested the retention of individual items. How are we justified in equating the two? One possibility is to assume that in the probe-digit experiment the subjects perceived and stored the digits as a series of overlapping pairs, rather than as single digits. In this case, the measure of interference would be given by the number of digit-pairs that follow any given pair, which is, of course, equal to the number of single digits that follow it. In the free-recall experiment, on the other hand, the subjects may have perceived the words as independent units, and the effective interference would then consist of single words, as we have in fact been assuming. The problem, then, can be restated as follows: why do pairs of digits and single words exert equal amounts of retroactive interference on like items in primary memory? There is little in the existing literature that sheds much light on this point.

Paired Associates

Our model should, of course, be able to describe ordered as well as free recall. We face serious problems, however, in attempting to apply it to serial learning: if a list is long enough to furnish a stable estimate of $S(i)$, the probability that a given item will be in PM at the time of testing is negli-

gible, since serial items are customarily tested in the order in which they were presented. We must therefore turn to paired associates. In a recent study, Tulving and Arbuckle (1963) systematically varied the positions of the items on the recall list, and we have therefore applied our hypothesis to their data in the manner described above.

Tulving and Arbuckle presented number-word pairs to their subjects and then tested for the recall of each word by presenting only the number with which it had been paired. They were interested in measuring probability of recall after one trial as a function of an item's serial position in both the original list and the test list. We have estimated $S(i)$ by averaging the recall probabilities for $i > 13$, excluding Items 1 and 2. The value of their serial position curve is fortunately constant in this region, as it was for free recall. Note that in this task, each pair presented after a given number and before the cue for its recall actually consists of *two* interfering items: a word plus a number. We have counted all items occurring between the test item and its recall—including the test number—as interfering items. We have analyzed the proportion of items presented in Positions 1 through 6 from the end of the stimulus list and tested in Positions 1 through 6 of the response list. These proportions are shown in Tulving and Arbuckle's Tables 2 and 4; we have pooled those that correspond to a given value of i . Thus i , or $n + m$ (where $n = j$ and $m = i - j$), ranges from 1 to 11. These data are presented in Figure 8, along with our own estimate of $P(i)$. Again, considering the variability of $S(i)$ that is not taken into account by our model, the fit between data and theory appears to be reasonably good.

In sum, then, we believe we can say

that the similarity between our probe-digit function and the various other, initially disparate, serial position curves shown in Figures 3-8 is consistent with the hypothesis that there is a primary memory store that is independent of any longer term store. The capacity of the primary store appears to be invariant under a wide variety of experimental conditions which do, however, affect the properties of the longer term store.

Single-Item Retention

Much of the experimental work on memory in the past 5 years has focused on measuring the retention of a single verbal item—or of a brief list of items—over short intervals. A widely used procedure which was introduced by Peterson and Peterson (1959) is to expose an item (for example, a meaningless three-letter sequence) to a subject; have him perform some task that presumably monopolizes his attention (such as counting backwards by three's) for a specified number of seconds; and, finally, at the end of this interval, have him attempt to recall the critical item. The universal finding has been that retention decreases monotonically with the length of the retention interval. It has generally been assumed that the subject does not rehearse during the retention interval, that a number spoken by him does not interfere with a trigram previously spoken by the experimenter, and that therefore the observed decline over time in the retention of such an item reflects the pure decay of its memory trace. This general conclusion is clearly inconsistent with our results, since we have found that the length of the retention interval as such—within the limits we tested, naturally—is of relatively little importance in determining retention loss.

In seeking for a way to account for

this discrepancy, it occurred to us to question the assumption that, in an experiment of the sort described above, the numbers spoken by the subject during the retention interval do not interfere with the memory trace of the item he is supposed to retain. Some experimenters have, after all, reported that dissimilar items seem to interfere with one another just as much as do similar ones in the immediate recall of very short lists (Brown, 1958; Pillsbury & Sylvester, 1940). What would happen, therefore, if we were to define a three-digit number uttered by a subject in the course of a simple arithmetic calculation—counting backwards—as one unit of mnemonic interference? Could our model then describe the forgetting of single items over brief intervals? We have attempted to fit the data of two experimenters, Loess (in press) and Murdock (1961), by converting the retention interval into a corresponding number of interfering items. Murdock's subjects were trained to count at a steady rate of one number per second, so the number of interfering items in his experiment is equal to the retention interval in seconds. Loess' subjects counted at a rate of one number every 1.5 seconds; we have therefore multiplied the length of his

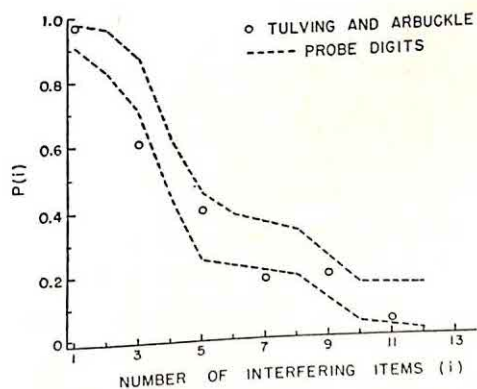


FIG. 8. Paired-associate data from Tulving and Arbuckle (1963) corrected for asymptote and response interference.

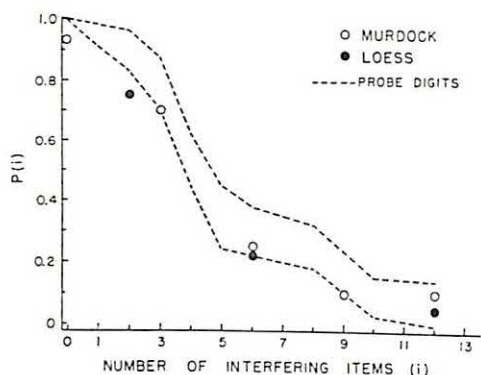


FIG. 9. The retention of three-item lists compared with the probe-digit function, (Loess' data denote the proportion of consonant trigrams recalled after various retention intervals; Murdock's data represent the average proportion of trigrams and word triads retained after a given interval.)

retention intervals by $\frac{2}{3}$ in order to obtain the equivalent number of interfering items. We estimate $S(i)$ in both cases by the relative frequency of recall at $i = 18$.⁶

The two sets of data, corrected for asymptote, are shown in Figure 9, along with the probe-digit function. The correspondence between them is reasonably close. It is possible, of course, that this agreement between theory and fact is simply a matter of luck, depending, as it does, on the arbitrary assumption that a three-digit number generated by the subject himself is psychologically equivalent to a one-digit number presented by the experimenter during the retention interval (as in the probe-digit study). Obviously we cannot draw any firm conclusions about the effect of interference on the retention of single items until this assumption is justified empirically. We can only point out that

⁶ We have also tried to analyze the results of Peterson and Peterson (1959) but without success. Part of the difficulty may result from the fact that their subjects may not have adhered strictly to a prescribed rate of counting during the retention interval (L. Peterson, personal communication, 1964).

the results of Murdock and Loess do not necessarily contradict our model.

DISCUSSION

We should at this point like to consider the general question of whether all verbal information is stored in the same system or whether, as we have assumed here, there are two independent mnemonic processes that contribute to retention even over very short intervals. The proponents of a unitary theory of memory, eloquently led by Melton (1963), have argued that recall after a few seconds is affected in very similar ways by the variables that govern recall over much longer intervals; and that therefore the distinction between a short-term memory mechanism, on the one hand, and a longer term mechanism, on the other, is purely arbitrary. The following facts have been cited in support of this argument:

1. Short-term retention improves, just as does long-term retention, when the material to be recalled is repeated before a test of retention, or when it is repeated between successive tests (Hebb, 1961; Hellyer, 1962).

2. Retention after a brief delay is subject to proactive interference, as is retention after a long delay (Keppel & Underwood, 1962; Loess, in press). Why, asks the unitary theorist, should we distinguish between short- and long-term retention if we cannot find any quantitative and experimentally manipulatable differences between them? This question might well be disturbing if one took the position that the two processes have sharply defined non-overlapping temporal boundaries such that items recalled within some critical interval after their initial occurrence must have been retrieved from one system, whereas items recalled beyond this interval must have been retrieved

from another. (Such a view would imply, interestingly enough, that an item would have to remain in a short-term storage for some specified number of seconds before passing into longer term storage, if it did so at all.)

But what if we do not require that the two systems be mutually exclusive? Then the probability that an item will be recalled will depend on both the probability that it is still in PM and the probability that it has entered into SM in the interval between its presentation and the start of the interfering sequence (or even during this sequence, if the subject is able to rehearse). All those variables that determine $S(i)$ for a given item—such as its position in a closely spaced series of tests, or the number of times it has been repeated—will then determine the observed proportion recalled after a brief interval. We believe we have shown, however, that $P(i)$ depends only on i and remains invariant with changes in $S(i)$; and we submit that most of the published data on short-term retention actually reflect the properties of both memory systems.

We would like to make one final point: the existence of some rather compelling introspective evidence in favor of two distinct mnemonic sys-

tems. PM, as we have defined it here, is best illustrated by a person's ability to recall verbatim the most recent few words in a sentence that he is hearing or speaking, even when he is barely paying attention to what is being said, or to what he is saying. Given that the flow of speech is intelligible, failures in the immediate recall of words we have just heard—errors of either omission, transposition, or substitution—are probably so rare as to be abnormal. Indeed, we believe that it would be impossible to understand or to generate a grammatical utterance if we lacked this rather remarkable mnemonic capacity. In order to recall a sentence verbatim at a later time, however, we usually have to rehearse it while it is still available in PM.

The same effect holds for meaningless arrangements of verbal items. If we present a subject with a random string of words, letters, or digits, and ask him to reproduce them in any order he chooses, he can maximize the number he recalls by "unloading" the last few items immediately. Most subjects in free-recall experiments report that these very late items tend to be lost if they are not recalled immediately, whereas items that came earlier in the list can be retrieved at leisure, if they

TABLE 3
PROPORTION OF ITEMS FREELY RECALLED AS A FUNCTION OF
STIMULUS INTERFERENCE AND PRESENTATION TIME

STIMULUS INTERFERENCE AND PRESENTATION TIME									
Number of intervening items	Seconds per item								
	Distributed					Massed			
	1	2	3	4	6	2	3	4	6
0	.96	.99	.97	1.00	.97	.98	.97	1.00	1.00
1	.82	.90	.91	.86	.89	.82	.96	.87	.91
2	.76	.81	.86	.82	.87	.75	.63	.76	.63
3	.54	.64	.73	.76	.65	.51	.58	.58	.50
4	.38	.40	.50	.57	.60	.40	.31	.51	.44
5	.21	.36	.36	.49	.48	.27	.36	.45	.44
5 + ^a	.14	.26	.32	.38	.45	.25	.31	.38	.42

^a Entries in this row represent the asymptotic value of $R(n)$.

can be recalled at all. In the colorful terminology of one such subject (Waugh, 1961), the most recent items in a verbal series reside temporarily in a kind of "echo box," from which they can be effortlessly parroted back. When an experienced subject is trying to memorize a list of serial items, moreover, he "fills up" successive echo boxes as the list is read to him and attempts to rehearse the contents of each. He will invariably lose some items if rehearsal is delayed too long or if he attempts to load his echo box with more items than it can hold. We think it very likely that the PM function describes the (variable) capacity of this mechanism. We would remind you in this connection that, within very broad limits, the rate at which someone is speaking does not affect your ability to follow his words—just as differences in the rate at which meaningless lists of digits are presented do not exert any profound effect on the PM function.

CONCLUSIONS

We have tried to demonstrate the existence of a short-term or PM system that is independent of any longer term or secondary store by showing that one function relating probability of recall to number of intervening items can describe a number of seemingly disparate sets of experimental results. In doing so, we have deliberately avoided discussing a number of problems raised in our analyses. Foremost in our list of problems is the definition of an item.

Certainly the idea of a discrete verbal unit is crucial to our theory. The interference effect that we have studied seems to be invariant over a broad class of units and combinations of units—single digits, nonsense trigrams, and meaningful words. How long a string of such primitive units can we combine and still have one item? Is an item determined by our grammatical habits? Is it determined by the duration of the verbal stimulus? Is it determined by both? We do not know.

We have also avoided discussing the possible rules whereby items now in PM are displaced by later items. Are items lost independently of one another, or do they hang and fall together? It may perhaps prove difficult to answer this question experimentally, but it should not be impossible.

Finally, at what stage in the processing of incoming information does our PM reside? Is it in the peripheral sensory mechanism? Probably not. The work of Sperling (1960) indicates that "sensory memory"—to use Peterson's (1963) phrase—decays within a matter of milliseconds, whereas we have dealt in our analysis with retention intervals on the order of seconds. Does storage in PM precede the attachment of meaning to discrete verbal stimuli? Must a verbal stimulus be transformed into an auditory image in order to be stored in PM, even if it was presented visually? We refer the reader to a recent paper by Sperling (1963) for some thoughts on the latter question.

APPENDIX

A formal discussion of the interaction between PM and SM can be provided by a simple three-state Markov process. The assumptions of the model are:

1. There are three states of memory: S , P , and the null state, G .

2. The probability of recalling an item from either State S or State P is unity: items cannot be recalled from the null state, but they may be guessed with Probability g .

3. Items can only pass into State S when they are rehearsed and, for the ex-

periments discussed in this paper, we assume that items are rehearsed only when they are presented. The probability that an item is stored in S , given that it was successfully rehearsed, is α .

4. Items in P are interfered with by later presentation of different items: the probability that an item returns to the null state on the presentation of the i th interfering item is δ_i .

The following equivalents hold between the terms defined for the Markov model and the terms defined in the body of the paper:

1. $P_i(S)$ is equivalent to $S(i)$.
2. $P_i(P)$ is equivalent to $P(i)[1 - S(i)]$.
3. δ_i is equivalent to $1 - P(i)$.

Now, define the random variable π with Value 1 if the test item is presented, and with Value 0 if some other (interfering) item is presented. (We can also let π be a probability—namely, the probability that the test item is presented. The formal statement of the model does not change with this redefinition.)

The transition probabilities for any given stimulus item (the test item) are specified by the matrix

$$\begin{array}{ccc} & S & P & G \\ \begin{array}{l} S \\ P \\ G \end{array} & \begin{bmatrix} 1 & 0 & 0 \\ \alpha\pi & (1-\pi)(1-\delta_i) + \pi(1-\alpha) & (1-\pi)\delta_i \\ \alpha\pi & (1-\alpha)\pi & 1-\pi \end{bmatrix} \end{array}$$

Unfortunately, it is difficult to work with transition matrices of this form (with time-varying parameters). One approximation would be to let $\delta_i = \delta$, independent of i . This approximation yields an exponential decay function of the form $P_i(P) = (1-\alpha)(1-\delta)^{i-1}$. This is clearly not correct for the results of our experiment (Figure 1); but, for some purposes, it may not be a bad approximation. A model very similar mathematically to that produced by this simple approximation for δ_i has been studied by Atkinson and Crothers (1964), who found it to be quite good for certain types of paired-associates experiments. Their model, however, is derived from quite different considerations.

For any experiments with controlled rehearsals, the probability that an item reaches State S (or SM) is completely independent of the properties of the short-term state (P or PM). This is true because, as far as State S is concerned, the general transition matrix can be reduced by combining States P and G to form the "lumped" State P' . The new matrix is

$$\begin{array}{cc} & S & P' \\ \begin{array}{l} S \\ P' \end{array} & \begin{bmatrix} 1 & 0 \\ \alpha\pi & 1-\alpha\pi \end{bmatrix} \end{array}$$

This is a simple one-element Markov model. This means that although the complete description of the verbal learning process requires a description of the short-term state, a study of only the long-term retention of items can ignore the short-term memory.

REFERENCES

- ATKINSON, R. C., & CROTHERS, E. J. A comparison of paired-associate learning models having different acquisition and retention axioms. *Journal of Mathematical Psychology*, 1964, 1, 285-315.
- BROADBENT, D. E. *Perception and communication*. New York: Pergamon Press, 1958.
- BROWN, J. Some tests of the decay theory of immediate memory. *Quarterly Journal of Experimental Psychology*, 1958, 10, 12-21.
- BUGELSKI, B. R. Presentation time, total time, and mediation in paired-associate learning. *Journal of Experimental Psychology*, 1962, 63, 409-412.
- CLARK, L. L., LANSFORD, T. E., & DALLENBACH, K. M. Repetition and associative learning. *American Journal of Psychology*, 1960, 73, 22-40.
- CONRAD, R. Decay theory of immediate memory. *Nature*, 1957, 179, 831-832.
- DEESE, J., & KAUFMAN, R. A. Sequential effects in recall of unorganized and sequentially organized material. *Journal of Experimental Psychology*, 1957, 54, 180-187.
- HEBB, D. O. *The organization of behavior*. New York: Wiley, 1949.
- HEBB, D. O. Distinctive features of learning in the higher animal. In J. F. Delafresnaye (Ed.), *Brain mechanisms and learning*. London: Oxford Univer. Press, 1961. Pp. 37-46.

- HELLYER, S. Supplementary report: Frequency of stimulus presentation and short-term decrement in recall. *Journal of Experimental Psychology*, 1962, 64, 650.
- JAMES, W. *The principles of psychology*. Vol. 1. New York: Holt, 1890. Ch. 16.
- KEPPEL, G., & UNDERWOOD, B. J. Proactive inhibition in short-term retention of single items. *Journal of Verbal Learning and Verbal Behavior*, 1962, 1, 153-161.
- LOESS, H. Proactive inhibition in short-term memory. *Journal of Verbal Learning and Verbal Behavior*, in press.
- MCGEOCH, J. A. Forgetting and the law of disuse. *Psychological Review*, 1932, 39, 352-370.
- MELTON, A. W. Implications of short-term memory for a general theory of memory. *Journal of Verbal Learning and Verbal Behavior*, 1963, 2, 1-21.
- MURDOCK, B. B., JR. The retention of individual items. *Journal of Experimental Psychology*, 1961, 62, 618-625.
- MURDOCK, B. B., JR. The serial position effect in free recall. *Journal of Experimental Psychology*, 1962, 64, 482-488.
- MURDOCK, B. B., JR. Interpolated recall in short-term memory. *Journal of Experimental Psychology*, 1963, 66, 525-532.
- PETERSON, L. R. Immediate memory: Data and theory. In C. N. Cofer, & Barbara Musgrave (Eds.), *Verbal behavior and learning: Problems and processes*. New York: McGraw-Hill, 1963. Pp. 336-353.
- PETERSON, L. R., & PETERSON, M. J. Short-term retention of individual verbal items. *Journal of Experimental Psychology*, 1959, 58, 193-198.
- PILLSBURY, W. B., & SYLVESTER, A. Retroactive and proactive inhibition in immediate memory. *Journal of Experimental Psychology*, 1940, 27, 532-545.
- POSNER, M. I. Immediate memory in sequential tasks. *Psychological Bulletin*, 1963, 60, 333-349.
- POSTMAN, L. The present status of interference theory. In C. N. Cofer (Ed.), *Verbal learning and verbal behavior*. New York: McGraw-Hill, 1961. Pp. 152-179.
- SPERLING, G. The information available in brief visual presentations. *Psychological Monographs*, 1960, 74 (11, Whole No. 498).
- SPERLING, G. A model for visual memory tasks. *Human Factors*, 1963, 5, 19-36.
- TULVING, E., & ARBUCKLE, T. Y. Sources of intratrial interference in immediate recall of paired associates. *Journal of Verbal Learning and Verbal Behavior*, 1963, 1, 321-334.
- UNDERWOOD, B. J., & KEPPEL, G. An evaluation of two problems of method in the study of retention. *American Journal of Psychology*, 1962, 75, 1-17.
- WAUGH, N. C. Free versus serial recall. *Journal of Experimental Psychology*, 1961, 62, 496-502.
- WAUGH, N. C. The effect of intralist repetition on free recall. *Journal of Verbal Learning and Verbal Behavior*, 1962, 1, 95-99.
- WAUGH, N. C. Immediate memory as a function of repetition. *Journal of Verbal Learning and Verbal Behavior*, 1963, 2, 107-112.

(Received March 10, 1964)

AN IRRITATIVE HYPOTHESIS CONCERNING THE HYPOTHALAMIC REGULATION OF FOOD INTAKE

ROBERT W. REYNOLDS

University of California, Santa Barbara

The experimental data relating to hypothalamic hyperphagia are re-examined in the light of a discussion of the possible effects of the lesion making procedure on the brain tissue. It is shown that, with the exception of 1 study, all of the data are compatible with the hypothesis that hyperphagia is a result of the chronic irritation of the ventrolateral hypothalamic "feeding center" by scar tissue surrounding electrolytic ventromedial hypothalamic lesions. This "irritative hypothesis" also accounts for the failure of the signs of hyperphagia to appear following ventromedial lesions produced with radio-frequency current. It is suggested that the general dual excitatory-inhibitory theory of the hypothalamic control of motivation and emotion needs reexamination.

One of the most intriguing phenomena in the repertoire of the physiological psychologist is the syndrome known as hypothalamic hyperphagia. It is readily obtainable by anyone with rudimentary skill in stereotaxic surgery. Its several signs suggest its relevance to the study of the physiological bases of motivation and emotion. These signs include postoperative voracity and obesity frequently accompanied by viciousness and genital atrophy. No attempt will be made here to carry through an exhaustive survey of the literature concerning this syndrome. Excellent surveys are already available in reviews by Anand (1961), Brobeck (1960), Ingram (1960), Larsson (1954), Mayer (1953), Rosenzweig (1962), Stellar (1954), and Teitelbaum (1961). The central experimental findings may, for the moment, be reduced to the following:

1. Bilateral electrolytic destruction of the region of the ventromedial hypothalamic nuclei is followed by a marked increase in food intake persisting for periods up to 40 or 50 days and leading to extreme obesity (Brobeck, Tepperman, & Long, 1943;

Hetherington & Ranson, 1940). This is known as hypothalamic hyperphagia.

2. Bilateral electrolytic destruction of the ventrolateral hypothalamic region is followed by aphagia and adipsia leading to death by inanition unless the animal is maintained by tube feeding (Anand & Brobeck, 1951; Anand, Dua, & Schoenberg, 1955; Morgane, 1961b; Teitelbaum, & Epstein, 1962).

3. Bilateral destruction of both the ventromedial and ventrolateral hypothalamic regions leads to aphagia (Anand & Brobeck, 1951).

4. Electrical stimulation of the ventrolateral region produces immediate eating behavior in satiated animals (Brügger, 1943; Larsson, 1954; Smith, 1956) as well as extended increases in food intake (Delgado & Anand, 1953).

5. Stimulation of the ventromedial hypothalamus produces a cessation of eating in food deprived animals (Anand & Dua, 1955; Morgane, 1961a; Smith, 1956).

6. Hypothalamic hyperphagic rats will not work as hard for food as normal rats (Miller, Bailey, & Stevenson, 1950; Teitelbaum, 1957). Presumably, therefore, they are not hyper-

motivated, but simply cannot achieve satiation due to the destruction of a "satiety center."

These data have generally led to a theory of the hypothalamic regulation of food intake in which activity in the ventrolateral hypothalamic "feeding center" is modulated by the inhibitory control of the ventromedial hypothalamic "satiety center." The overeating and obesity in hypothalamic hyperphagia are normally explained, therefore, as the result of the animal's inability to satiate itself following removal of its "satiety center." The ventromedial hypothalamic nuclei, therefore, are presumed to function in the regulation of food intake as inhibitory centers, removal of which leads to "release" eating.

This theory is extremely appealing in its simplicity and comprehensiveness. It is possible, however, that it is based on an artifact of the lesion making procedure. In studies involving destruction of brain tissue there is generally an implicit, basic assumption that the behavioral results must be due to the removal of the tissue destroyed in the lesioning process. It would seem desirable, however, to call even this assumption into question, since it is quite possible that there may be some residual scar tissue which could serve as a source of irritation to the surrounding tissue. Specifically, we should ask precisely what the lesioning process does to the tissue of the central nervous system. The procedure almost universally used to produce lesions in studies of the physiological basis of motivation and emotion is electrolysis. Direct current, normally anodal, is passed through a monopolar electrode which is insulated except at the tip. The uninsulated tip is placed in the neural structure to be destroyed. The circuit is normally completed either by a large cathode inserted into the rec-

tum, or by connecting the metal frame of the stereotaxic instrument to the negative side of the power supply. The tissue surface in contact with the cathode is sufficiently large that the current density is very small, and no damage is done at that point. The current density at the anodal electrode in the brain is, conversely, very high because of the very small surface involved.

The process that occurs at the tip of the anode when current passes depends on the nature of the surrounding medium and the material out of which the anode is made. It is probably safe to assume for our purposes that the brain tissue acts essentially as a saline bath, i.e., it is a reasonably good electrolyte. Given this, the type of reaction that will occur at the anode depends on the electrode potential of the electrode material. Metals with an electrode potential less positive than hydroxyl ion (i.e., less than $+0.40$ for half cell reduction reaction) will go into solution as the corresponding metallic ions. Thus, anodes made of copper, nickel, iron, and chromium will deposit metallic ions in brain tissue. According to the investigations of MacIntyre, Bidder, and Rowland (1959), and Rowland, MacIntyre, and Bidder (1960), the destruction of brain tissues with such electrodes is a function of the invasion of the tissue by these metallic ions. In addition to these metals, platinum is frequently used as the electrode material for making lesions. However, since platinum has a more positive electrode potential than hydroxyl ion ($+0.863$), the hydroxyl ion reaction is favored. Therefore, cytolysis in this case is not caused by the depositing of platinum ions in the tissue, but, rather, is presumably the mechanical result of the generation of oxygen gas bubbles at the electrode tip. The lesions produced when the

current is reversed (i.e., the electrode is made the cathode) are negligible with steel, copper, or nickel electrodes (Rowland et al., 1960). However, a platinum electrode is apparently even more effective (Horsley & Clarke, 1908) when used as the cathode, presumably because for a given amount of current, the volume of gas generated is twice as much at the cathode (hydrogen) as at the anode (oxygen). It is clear that either the introduction of metallic ions or mechanical deformation by the generation of gas bubbles will destroy the cells around the tip of the electrode. This is not all that happens, however. The introduction of metallic ions in nerve tissue evokes a glial reaction and generally results in the formation of a scar. When the tissue at the electrode tip is highly vascularized, as is the case in the hypothalamus, there may be rupture of blood vessels followed by hemorrhage

and the formation of scar tissue. The formation of hemorrhagic scar tissue is even more likely following the mechanical destruction of highly vascularized tissue by a platinum anode. It is highly probable, therefore, that in addition to destroying nerve cells, electrolytic lesions will leave some scar tissue around the area of destruction. It is a well-known clinical observation (Walker, 1949) that scar tissue may serve as an irritative focus for the chronic production of epileptic seizures or the signs of basal ganglia disease (Glees, 1961, p. 308). It would seem reasonable, then, to hypothesize that electrolytic lesions, particularly in highly vascularized regions, may have chronic irritative effects on surrounding tissue. It is a hypothesis, however, that deserves direct experimental verification.

The preceding analysis was necessitated by the observation (Reynolds,

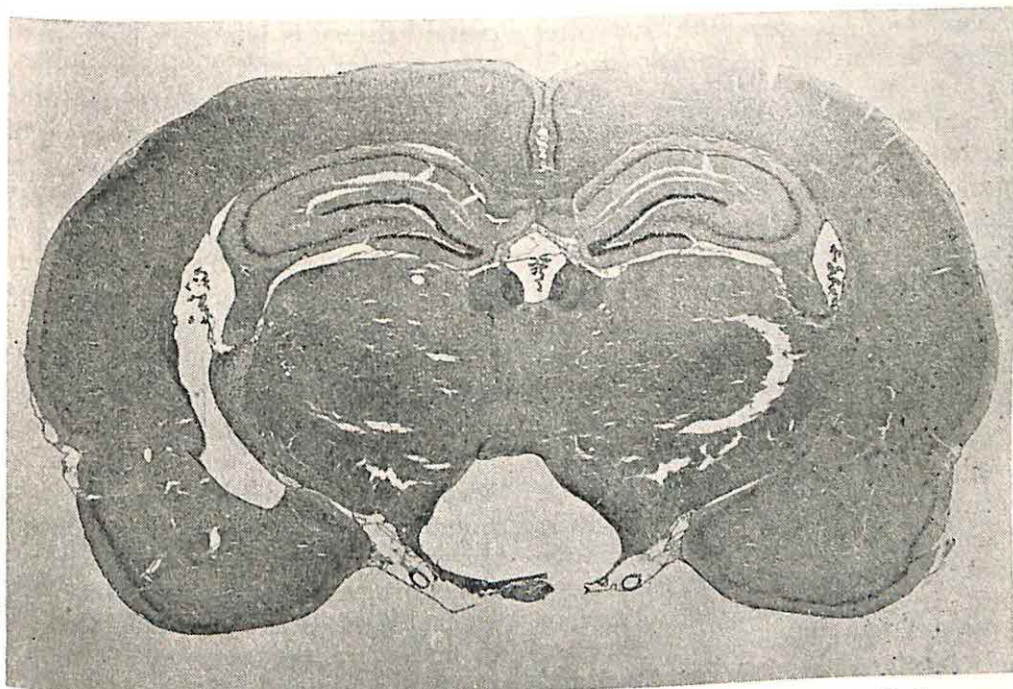


FIG. 1. Photomicrograph of coronal section through greatest extent of radio-frequency lesion in ventromedial region of rat hypothalamus. (The only discernable chronic effect of this lesion was diabetes insipidus.)

1963c) that destruction of the ventromedial hypothalamic nuclei by radio-frequency thermocoagulation did not produce hyperphagia. The radio-frequency lesions in one of the animals in that study are shown in Figure 1. These lesions completely destroyed the medial part of the hypothalamus including the ventromedial nuclei without producing the chronic signs of hyperphagia following similarly placed electrolytic lesions. The lesions were bounded laterally by the fornices and extended from the plane of the posterior end of the paraventricular nucleus to the anterior end of the mamillary bodies. It should be noted that this was not simply ventricular distention, since there was no evidence of residual ventricular ependyma around the lesion or tissue compression lateral to the lesion. This original study was subsequently confirmed in a replication involving a direct comparison of 55 rats with electrolytic lesions and 40 rats with radio-frequency lesions of comparable size in the ventromedial hypothalamus. Averill and Purves (1963) also failed to produce hyperphagia with radio-frequency lesions in the hypothalamus, although they attribute this failure to the small size of their lesions.

Radio-frequency lesions are produced by a surgical procedure identical to that used for electrolytic lesions except that the current passed through the electrode is a biphasic sine-wave current at a frequency of 2 megacycles per second. Because of the biphasic character of this current, there is no residual deposit of metallic ions in the tissue. Cytolysis is produced through heat at the electrode tip (Aronow, 1960). In addition to destroying nerve cells, the heat can cauterize small blood vessels in the region, and thus tend to minimize the incidence of hemorrhage. The formation of scar

tissue, therefore, should be relatively infrequent in radio-frequency lesions. The microscopic appearance of the lesions in general confirms this.

Since removal of the ventromedial hypothalamic nuclei by radio-frequency lesions does not produce a chronic change in food intake, the hypothalamic hyperphagia resulting from electrolytic destruction of the ventromedial nuclei cannot simply be attributed to removal of a ventromedial hypothalamic "satiety center." We must, therefore, find an alternative explanation. The position outlined above, that scar tissue produced by electrolytic lesions may have an irritative effect on the surrounding tissue, may be adopted as a working hypothesis. Hyperphagia may arise, therefore, from the chronic irritation of the adjacent ventrolateral "feeding center." This hypothesis, however, is dependent on the theory that the ventrolateral hypothalamus is indeed a "feeding center." Since the "feeding center" theory is based, in large part, on data from electrolytic lesion studies, it seemed necessary to ascertain the effect of radio-frequency destruction of the ventrolateral hypothalamus as a first step in exploring the "irritative" hypothesis. No difference was found in the effectiveness of radio-frequency and electrolytic lesions in producing aphagia (Reynolds, 1963b).

It should be noted that transient irritative effects may be produced by massive radio-frequency lesions which have no long-term, chronic irritative effects simply as a function of surgical trauma and local tissue edema. It is the chronic effects, however, which are of importance in the present discussion. It should also be noted that the use of radio-frequency current does not guarantee that the lesions will be free of scar tissue or irritative effects. It simply minimizes the probability of their occurrence. Conversely, the use

of electrolysis does not guarantee that the lesions will have irritative effects, but rather it maximizes the probability thereof. This would presumably account for the fact that on occasion electrolytic destruction of the ventromedial hypothalamus also fails to produce hyperphagia.

At the beginning of this paper there were listed six main points regarding the hypothalamic regulation of food intake. The first point, that electrolytic ventromedial lesions produce hyperphagia, has been accounted for already by the "irritative" hypothesis itself. It is also instructive to note that Kennedy (1950) finds

that the most effective lesions lie ventrolateral to the ventromedial nuclei, and often appear to cause little direct damage to the nuclei themselves.

Furthermore, the first implication of the tuberal hypothalamic region in the control of food intake was quite likely the description of Fröhlich's Syndrome. One of the signs of this syndrome is an extreme obesity. The underlying pathology is an expanding tumor in the infundibular region. It does not seem unreasonable to assume that such tumors could have irritative effects. Points 2, 3, and 4, concerning the ventrolateral "feeding" center, are compatible with the "irritative" hypothesis—indeed, they would be predicted by it. But what of the remaining two points?

On a priori grounds, it may be argued that Point 5—that stimulation of the ventromedial hypothalamus produces a cessation of eating in hungry animals—has meaning only in the context of the "satiety center" theory. Indeed, Krasne (1962) has recently shown that stimulation of the ventromedial hypothalamus at an intensity capable of producing a cessation of eating would also prevent drinking and

could be used as the motivational stimulus to produce escape behavior. Krasne concludes, therefore,

the stopping of eating caused by stimulation of the ventromedial hypothalamus and adjacent areas may not represent a primary effect on hunger,

but may instead be a result of a generalized "upsetting or distracting" effect of the stimulation.

The studies purporting to show that hyperphagics are not overmotivated cited in Point 6 should be reexamined. In the Miller, Bailey, and Stevenson (1950) study animals with hypothalamic lesions were shown to work less for food than normals on several different tasks. The significance of this study must be discounted, however, since there is no convincing evidence presented that any of the animals were in fact hyperphagic. The authors state,

Placed on a diet of dry, ground lab chow, four of the operated animals (to be called Operated High-Eaters) ate 30%–85% more than the average of the controls; seven (Operated Low-Eaters) are approximately the same or less than the controls [p. 257].

It is quite likely, however, that 5 of the 10 control animals also ate more than the average of the controls, but in the absence of any information about the range of food intake in the control group there is no way to determine the degree to which there was any overlap with the "Operated High-Eaters." The authors do not indicate that there was a statistically significant difference between the Operated High-Eaters and the controls. In view of their reference to a statistically significant difference with respect to another comparison in the next sentence, one is left with the impression that there was not. A superficial reading, however, might give the opposite impression. Furthermore, a

post facto classification into groups (Operated High-Eaters and Operated Low-Eaters) on the basis of the dependent variable is hardly a legitimate procedure. Perhaps more important, however, are the weight data as given in Figure 1. It appears that the Operated High-Eaters had an average weight of about 370 grams at the time of operation, while the controls averaged about 345 grams. Thus the Operated High-Eaters started off weighing about 25 grams more than the controls. Forty days after the operations the average weight of the Operated High-Eaters was about 425 grams and the average weight of the controls was about 400 grams. Again a difference of about 25 grams. Thus, the controls gained weight at almost exactly the same rate as the alleged hyperphagics. This rate was about 1.5 gram per day, which certainly cannot be considered hyperphagia. It is true that when they were put on a high fat diet both operated groups gained weight faster than the controls. But this means that their "hyperphagia" appeared only with a high fat diet, which is contrary to normal experience with hyperphagia. This study is further weakened by a footnote in a subsequent paper by Miller (1955) which states that, "under somewhat different conditions, which are not yet well understood by us, the rats with hypothalamic lesions can not only eat more but also work harder for food [p. 141]."

Teitelbaum's (1957) paper also contains some methodological problems. He compared bar-press rates for pellets with dynamic hyperphagics, obese hyperphagics, and normal controls on fixed ratio schedules ranging from 1 to 256. Animals were given access to the bar for 12 hours each day, with no food available for the other 12 hours. After 2 or 3 days of stabilization on

continuous reinforcement, the reinforcement ratio was increased through the Steps 1, 4, 16, 64, and 256 with 2 days at each step. For the first 3 steps, the dynamic hyperphagics actually gave a higher average number of responses than the normal animals. At ratios of 64 and 256:1 the response output of the hyperphagics dropped off, while that of the normals continued to increase. It should be noted, however, that even the normals gave only a mean of about 4,500 responses per 12-hour period at the 256:1 ratio. A well-trained animal will approximate that many responses in 1 hour at such a high ratio. It is apparent, therefore, that while this procedure does not exactly constitute extinction, it also does not represent "typical" performance at these ratios. The exact sense in which this is a fair test of relative motivation is certainly not clear. There is the further confounding, noted by Teitelbaum himself, that the hyperphagics are very low-activity animals, and do not show the increase in activity as a function of food deprivation which is characteristic of normals. Hamilton (1963) has suggested that the failure of the hyperphagic to increase its work rate at the higher ratios may be due to an increased tendency to hyperthermia which is aggravated at the higher activity levels required by high ratios. More recently, Hamilton and Brobeck (1964) replicated some aspects of the Teitelbaum (1957) and Hamilton (1963) studies using hyperphagic monkeys. They showed that

the hyperphagic monkey works for more food than controls at all FR schedules up to ratios where extinction in all Ss makes any difference meaningless [p. 278].

Falk (1961) has shown a very high curvilinear correlation (.84) between increase in response rate on a variable interval schedule (VI 1') following ventromedial lesions and weight at-

tained after surgery and the post-surgery testing. This would suggest that the hyperphagics, at least in the initial stages, will in fact work harder than normals. Teitelbaum (1961) has commented that during the postsurgical testing in Falk's experiment the animals were maintained in a deprived or semistarved condition and that this may have introduced a new factor. It should be pointed out, however, that the obese hyperphagics in Teitelbaum's own experiment apparently received on the average only 1 gram of food altogether over the last 4 days of testing, and the dynamics about 5 grams. Considering that the dynamics weighed only 20 grams more than the controls at the start of the testing it is quite likely that they would have lost the 70 grams necessary to reduce their weights to 80% of the control weight during these 4 days of relative deprivation. Finally, Balinska (1963) has shown an increase over normal in the performance of a conditioned-reflex Type II in dynamic hyperphagic rabbits. In any case, it is probably fair to conclude that the notion that hyperphagics are not overmotivated, but rather are simply insatiable, is not very firmly established.

There are other studies bearing on the role of ventromedial hypothalamus in hunger motivation that should be mentioned. Anand, Dua, and Singh (1961), and Anand, Chhina, and Singh (1962) have studied the effect of systemic glucose and insulin injections on activity recorded from the ventromedial and ventrolateral hypothalamic regions through both macro- and micro-electrodes. They report an increase in activity in the ventromedial region following glucose and a decrease following insulin injections. The effects on the ventrolateral region were the reverse. These results support Mayer's (1953) "glucostatic" theory

in which the ventromedial hypothalamic "satiety center" is activated by high blood glucose levels or rates of glucose utilization. Sutin (1963), however, reports no effect of glucose injections on potentials evoked in the ventromedial nucleus by stimulation of the amygdala and septal region. Although both Sutin's and Anand's groups used both anesthetized and unanesthetized animals, the resolution of these conflicting results is made difficult by Brooks' (1959) observations on the very marked influence of depth of anesthesia on activity in the ventromedial hypothalamus. At best, we can only say that the electrophysiological evidence is equivocal. Recently this point has been subjected to behavioral examination. If the ventromedial nucleus were indeed the "glucostat," glucose injections should not depress eating in animals with ventromedial lesions. Russek and Morgane (1963), however, found glucose suppression of food intake in two cats made hyperphagic by electrolytic lesions in the ventromedial hypothalamus to be equivalent to that found in normal control animals. Reynolds and Kimm, moreover, found an exaggerated anorexic effect of glucose injections in hypothalamic hyperphagic rats in comparison to normal controls. This is reminiscent of the similar "paradoxical" results obtained by Epstein (1959), Stowe and Miller (1957), and Reynolds (1959) on the effect of amphetamine on food intake in hyperphagic rats. These studies tested the hypothesis that amphetamine has its anorexic effect by stimulating the ventromedial hypothalamic "satiety center." It was found, however, that there was an even greater effect in the absence of the supposed target. These results still have not been satisfactorily explained.

Hoebel and Teitelbaum (1962) have combined the techniques of electrolytic

ablation, intrahypothalamic chemical injection, and intracranial self-stimulation in an effort to elucidate the motivational roles of the ventromedial and ventrolateral hypothalamus. Their results generally seem to support the theory that the ventromedial hypothalamus normally has an inhibitory effect on the ventrolateral hypothalamus. This appears to be a fairly elaborate and comprehensive study. However, the conclusions drawn by the authors themselves are quite shattering. Their last two sentences state,

When an animal is hungry, lateral hypothalamic self-stimulation is more reinforcing; when satiated, it is less so. It may be that the pleasure of lateral hypothalamic self-stimulation is similar to the gratification obtained by eating [p. 376].

We may ignore the hedonic overtones. The essential point is that stimulation of the ventrolateral hypothalamus is positively reinforcing. It is more reinforcing when the animal is hungry. This, however, is the "feeding center," activation of which leads to eating. Under normal physiological conditions, high levels of activity in the "feeding center" are supposed to correspond to a condition of hunger. If ventrolateral self-stimulation has, as the authors suggest, any relationship to normal physiological activation, the feeding-center theory should lead us to believe that such self-stimulation must be related to the induction of hunger rather than the "gratification obtained by eating." This in turn leads us to the perplexing conclusion that the induction of hunger is positively reinforcing, or that animals should learn responses which result in their becoming hungry. This may or may not be true, but it certainly a novel inversion of most current theories of hunger motivation. Whatever the implications of this study for the "irritative hypothesis," it would be potentially quite devastating for the

feeding-center theory. The impact of this study is further weakened, however, by Morgane's (1962) demonstration of the dissociation of the self-stimulation and "feeding" systems. Medial fore-brain bundle lesions anterior and posterior to the ventrolateral hypothalamic feeding center "significantly suppressed" the self-stimulation response in the ventrolateral hypothalamus without affecting the feeding response elicited by stimulation through the same electrode.

Using bilateral cannulae implanted in the ventromedial or ventrolateral hypothalamus, Epstein (1960) observed the effects on eating behavior of local injections of procaine and hypertonic saline. His findings were in line with the "satiety-feeding center" theory, i.e., ventromedial injections of saline depressed food intake in deprived rats while procaine injections induced eating in satiated animals. The effects were reversed for ventrolateral injections. All of these results can be comprehended by the irritative hypothesis except for the effect of procaine on the ventromedial hypothalamus. It is assumed that procaine has an inhibitory effect on hypothalamic cells. The ventromedial injections, therefore, supposedly represent a temporary ablation of the "satiety center." Epstein's behavioral data are impressive, and are clearly incompatible with the irritative hypothesis. However, they are also at odds with my observation that radio-frequency ablation of the ventromedial hypothalamus does not produce hyperphagia. It is a discrepancy that deserves empirical resolution.

In a study just completed, electrolytic lesions were placed in rats using stereotaxic coordinates appropriate for the region of the Fields of Forel, just dorsal to the ventrolateral hypothalamic region. Although there were no major

long-term changes in body weight with the exception of three cases of aphagia, there was an extremely significant increase in the chewing activity of these animals. To obtain a rough estimate of this chewing activity, periodic measures were taken of the amount of food crumbs accumulating under the cages over 24-hour periods. The animals with the electrolytic lesions averaged over 20 grams of food crumbs per 24-hour period in comparison with 5 grams for the normal controls. It should be noted that these animals were not ingesting increased quantities of food, since they were not gaining appreciably in body weight. It could be argued that the Fields of Forel are inhibitory centers which normally function to suppress chewing behavior. It seems much more likely, however, to suppose that these lesions were having a chronic irritative effect on the "far lateral" part of the feeding center, which Morgane (1961a) believes to control the motor and reflexive aspects of eating behavior. The fact that hyperphagia did not appear in these animals may be attributed to the fact that the lesions were not close enough to the motivational component, which, according to Morgane, lies more medially in the medial forebrain bundle, just lateral to the ventromedial region. Any attempt to approach the motivational component more closely from a lateral or dorsal direction would involve destruction of the motor control component, thus preventing the appearance of hyperphagia because of the inability of the animals to chew. Indeed, Williams and Teitelbaum (1959) found that although animals with such lesions were aphagic when maintained on a diet of dry pellets and water, they became hyperphagic when placed on a liquid diet which presumably did not require chewing. A reasonable explanation for these data in terms of the

"release phenomenon" type of theory is not readily apparent. They can be accounted for without difficulty, however, in the context of the "irritative" hypothesis.

The irritative hypothesis obviously does not provide an alternative theory concerning the hypothalamic regulation of food intake. It does suggest an alternative explanation to the satiety-center theory for the experimental phenomenon of hypothalamic hyperphagia. At the very least, the above discussion indicates the need for a re-examination of the problem of the hypothalamic control of food intake. Analogously, we may question the general excitatory-inhibitory dual control theory (Stellar, 1954, 1960) of motivation. It may be that the "release" phenomena so widely observed following electrolytic destruction of postulated "inhibitory" centers are the result of irritation rather than ablation. Consistent with this position is the finding (Reynolds, 1963a) that, as in hypothalamic hyperphagia, the signs of pulmonary edema fail to appear when the lesions are made by radio-frequency thermocoagulation. This syndrome has been previously attributed to the release of an edemagenic center from inhibition following electrolytic lesions (Maire & Patton, 1956). Law and Meagher (1958) have demonstrated the elimination of sexual behavior following medial hypothalamic lesions while hypersexuality followed electrolytic lesions either anterior or posterior to this region. The postulation of two inhibitory centers, one anterior and one posterior to the excitatory center, would seem less parsimonious in the absence of other evidence than the suggestion that the anterior and posterior lesions are irritating the excitatory center which lies between them. Indeed, Hillarp, Olivecrona, and Silferskiöld (1954) have suggested that the

hypersexuality they observed immediately following preoptic lesions may be an irritative rather than a release effect. Everett (1961) has similarly reported that the production of ovulation in rats under pentobarbital anesthesia was consistently obtained with small electrolytic lesions in the preoptic region, while "control lesions of comparable dimensions produced by high-frequency electrocautery have been ineffective [p. 111]." Everett suggested that the positive effects obtained with the electrolytic lesions were the result of irritation. It is instructive to note that Nakao (1958), Glusman and Roizin (1960), and Wasman and Flynn (1962) find that stimulation of regions around the ventromedial hypothalamus elicits vicious behavior very similar to that which frequently appears chronically following ventromedial electrolytic lesions (Ingram, Knot, Wheatley, & Summers, 1951; Wheatley, 1944). The proximity of the points in the hypothalamus, stimulation of which yields gastric ulceration (Feldman, Behar, & Birnbaum, 1961; Long, Leonard, Story, & French, 1962), to the regions in which electrolytic lesions produce the same effect (Long, Leonard, Chou, & French, 1962) should also be grounds for suspicion.

The hypothalamic regulation of motivation and emotion has been an important integrating formulation for physiological psychology. The above discussion however, demonstrates the necessity for a reevaluation. If the theoretical structure rests on an artifact we should be prepared to salvage what we can, but also to begin work on a new and sounder edifice.

REFERENCES

- ANAND, B. K. Nervous regulation of food intake. *Physiological Review*, 1961, 41, 677-708.
- ANAND, B. K., & BROBECK, J. R. Hypothalamic control of food intake in rats and cats. *Yale Journal of Biology and Medicine*, 1951, 24, 123-140.
- ANAND, B. K., CHHINA, G. S., & SINGH, B. Effect of glucose on the activity of hypothalamic "feeding centers." *Science*, 1962, 138, 597-598.
- ANAND, B. K., & DUA, S. Feeding responses induced by electrical stimulation of the hypothalamus in cats. *Indian Journal of Medical Research*, 1955, 43, 113-122.
- ANAND, B. K., DUA, S., & SHOENBERG, K. Hypothalamic control of food intake in cats and monkeys. *Journal of Physiology*, 1955, 127, 143-152.
- ANAND, B. K., DUA, S., & SINGH, B. Electrical activity of the hypothalamic 'feeding centres' under the effect of changes in blood chemistry. *Electroencephalography and Clinical Neurophysiology*, 1961, 13, 54-59.
- ARONOW, S. The use of radio-frequency power in making lesions in the brain. *Journal of Neurosurgery*, 1960, 17, 431-438.
- AVERILL, R. L. W., & PURVES, H. D. Differential effects of permanent hypothalamic lesions on reproduction and lactation in rats. *Journal of Endocrinology*, 1963, 26, 463-477.
- BALINSKA, H. Food preference and conditioned reflex type II activity in dynamic hyperphagic rabbits. *Acta Biologicae Experimentalis, Warsaw*, 1963, 23, 33-44.
- BROBECK, J. R. Regulation of feeding and drinking. In J. Field (Ed.), *Handbook of physiology. Section 1: Neurophysiology*. Vol. 2. Baltimore: Williams & Wilkins, 1960. Pp. 1197-1206.
- BROBECK, J. R., TEPPERMAN, J., & LONG, C. N. H. Experimental hypothalamic hyperphagia in the albino rat. *Yale Journal of Biology and Medicine*, 1943, 15, 831-853.
- BROOKS, D. C. Electrical activity of the ventromedial nucleus of the hypothalamus. *American Journal of Physiology*, 1959, 197, 829-834.
- BRÜGGER, M. Fresstrieb als hypothalamisches Symptom. *Helvetica physiologica et pharmacologica acta*, 1943, 1, 183-198.
- DELGADO, J. M. R., & ANAND, B. K. Increased food intake induced by electrical stimulation of the lateral hypothalamus. *American Journal of Physiology*, 1953, 172, 162-168.
- EPSTEIN, A. N. Suppression of eating and drinking by amphetamine and other drugs in normal and hyperphagic rats. *Journal*

- of *Comparative and Physiological Psychology*, 1959, 52, 37-45.
- EPSTEIN, A. N. Reciprocal changes in feeding behavior produced by intrahypothalamic chemical injections. *American Journal of Physiology*, 1960, 199, 969-974.
- EVERETT, J. W. The preoptic region of the brain and its relation to ovulation. In C. A. Villee (Ed.), *Control of ovulation*. New York: Pergamon Press, 1961. Pp. 101-112.
- FALK, J. L. Comments on Dr. Teitelbaum's paper. In M. R. Jones (Ed.), *Nebraska symposium on motivation: 1961*. Lincoln: Univer. Nebraska Press, 1961. Pp. 65-68.
- FELDMAN, S., BEHAR, A. J., & BIRNBAUM, D. Gastric lesions following hypothalamic stimulation. *Archives of Neurology*, 1961, 4, 308-317.
- GLEES, P. *Experimental neurology*. London: Oxford Univer. Press, 1961.
- GLUSMAN, M., & ROIZIN, L. Role of the hypothalamus in the organization of agonistic behavior in the cat. In M. D. Yahr (Ed.), *Transactions of the American Neurological Association*, 1960. New York: Springer, 1960. Pp. 177-181.
- HAMILTON, C. L. Interactions of food intake and temperature regulation in the rat. *Journal of Comparative and Physiological Psychology*, 1963, 56, 476-488.
- HAMILTON, C. L., & BROBECK, J. R. Hypothalamic hyperphagia in the monkey. *Journal of Comparative and Physiological Psychology*, 1964, 57, 271-278.
- HETHERINGTON, A. W., & RANSON, S. W. Hypothalamic lesions and adiposity in the rat. *Anatomical Record*, 1940, 78, 149-172.
- HILLARP, N. A., OLIVECRONA, H., & SILFERSKIÖLD, W. Evidence for the participation of the preoptic area in male mating behavior. *Experientia*, 1954, 10, 224-225.
- HOEBEL, B. G., & TEITELBAUM, P. Hypothalamic control of feeding and self-stimulation. *Science*, 1962, 135, 375-377.
- HORSLEY, V., & CLARKE, R. H. The structure and functions of the cerebellum examined by a new method. *Brain*, 1908, 31, 45-124.
- INGRAM, W. R. Central autonomic mechanisms. In J. Field (Ed.), *Handbook of physiology. Section 1: Neurophysiology*. Vol. 2. Baltimore: Williams & Wilkins, 1960. Pp. 951-978.
- INGRAM, W. R., KNOTT, J. R., WHEATLEY, M. D., & SUMMERS, T. D. Physiological relationships between hypothalamus and cerebral cortex. *Electroencephalography and Clinical Neurophysiology*, 1951, 3, 37-58.
- KENNEDY, G. C. The hypothalamic control of food intake in rats. *Proceedings of the Royal Society (Series B)* 1950, 137, 535-549.
- KRASNE, F. B. General disruption resulting from electrical stimulation of ventromedial hypothalamus. *Science*, 1962, 138, 822-823.
- LARSSON, S. On the hypothalamic organization of the nervous mechanism regulating food intake. Part I. Hyperphagia from stimulation of the hypothalamus and medulla in sheep and goats. *Acta Physiologica Scandinavica*, 1954, 32, Suppl. 115, 1-40.
- LAW, T., & MEAGHER, W. Hypothalamic lesions and sexual behavior in the female rat. *Science*, 1958, 128, 1626-1627.
- LONG, D. M., LEONARD, A. S., CHOU, S. N., & FRENCH, L. A. Hypothalamus and gastric ulceration: I. Gastric effects of hypothalamic lesions. *Archives of Neurology*, 1962, 7, 167-175.
- LONG, D. M., LEONARD, A. S., STORY, J., & FRENCH, L. A. Hypothalamus and gastric ulceration: II. Production of gastrointestinal ulceration by chronic hypothalamic stimulation. *Archives of Neurology*, 1962, 7, 176-183.
- MACINTYRE, W. J., BIDDER, T. G., & ROWLAND, V. The production of brain lesions with electric currents. In H. Quastler & H. J. Morowitz (Eds.), *Proceedings of the first national biophysical conference*. New Haven: Yale Univer. Press, 1959. Pp. 723-732.
- MAIRE, F. W., & PATTON, H. D. Neural structures involved in the genesis of "pre-optic pulmonary edema," gastric erosions and behavior changes. *American Journal of Physiology*, 1956, 184, 345-350.
- MAYER, J. Genetic, traumatic and environmental factors in the etiology of obesity. *Physiological Review*, 1953, 33, 472-508.
- MILLER, N. E. Shortcomings of food consumption as a measure of hunger: Results from other behavioral techniques. *Annals of the New York Academy of Sciences*, 1955, 63, 141-143.
- MILLER, N. E., BAILEY, C. J., & STEVENSON, J. A. F. Decreased "hunger" but increased food intake resulting from hypothalamic lesions. *Science*, 1950, 112, 256-259.
- MORGANE, P. J. Electrophysiological studies of feeding and satiety centers in the rat. *American Journal of Physiology*, 1961, 201, 838-844. (a)

- MORGANE, P. J. Medial forebrain bundle and "feeding centers" of the hypothalamus. *Journal of Comparative Neurology*, 1961, 117, 1-26. (b)
- MORGANE, P. J. Dissociation of hypothalamic self-stimulation and primary "feeding" systems. *Clinical Research*, 1962, 10, 185.
- NAKAO, H. Emotional behavior produced by hypothalamic stimulation. *American Journal of Physiology*, 1958, 194, 411-418.
- REYNOLDS, R. W. The effect of amphetamine on food intake in normal and hypothalamic hyperphagic rats. *Journal of Comparative and Physiological Psychology*, 1959, 52, 682-684.
- REYNOLDS, R. W. Pulmonary edema as a consequence of hypothalamic lesions in rats. *Science*, 1963, 141, 930-932. (a)
- REYNOLDS, R. W. Radio frequency lesions in the ventrolateral hypothalamic "feeding center." *Journal of Comparative and Physiological Psychology*, 1963, 56, 965-967. (b)
- REYNOLDS, R. W. Ventromedial hypothalamic lesions without hyperphagia. *American Journal of Physiology*, 1963, 204, 60-62. (c)
- ROSENZWEIG, M. R. The mechanisms of hunger and thirst. In L. Postman (Ed.), *Psychology in the making*. New York: Knopf, 1962. Pp. 73-143.
- ROWLAND, V., MACINTYRE, W. J., & BIDDER, T. G. The production of brain lesions with electric currents: II. Bidirectional currents. *Journal of Neurosurgery*, 1960, 17, 55-69.
- RUSSEK, M., & MORGANE, P. J. Anorexic effect of intraperitoneal glucose in the hypothalamic hyperphagic cat. *Nature*, 1963, 199, 1004-1005.
- SMITH, O. A. Stimulation of lateral and medial hypothalamus and food intake in the rat. *Anatomical Record*, 1956, 124, 363-364.
- STELLAR, E. The physiology of motivation. *Psychological Review*, 1954, 61, 5-22.
- STELLAR, E. Drive and motivation. In J. Field (Ed.), *Handbook of physiology. Section I: Neurophysiology*. Vol. 3. Baltimore: Williams & Wilkins, 1960. Pp. 1501-1528.
- STOWE, F. R., & MILLER, A. T. The effect of amphetamine on food intake in rats with hypothalamic hyperphagia. *Experientia*, 1957, 13, 114-115.
- SUTIN, J. An electrophysiological study of the hypothalamic ventromedial nucleus in the cat. *Electroencephalography and Clinical Neurophysiology*, 1963, 15, 786-795.
- TEITELBAUM, P. Random and food-directed activity in hyperphagic and normal rats. *Journal of Comparative and Physiological Psychology*, 1957, 50, 486-491.
- TEITELBAUM, P. Disturbances in feeding and drinking behavior after hypothalamic lesions. In M. R. Jones (Ed.), *Nebraska symposium on motivation: 1961*. Lincoln: Univer. Nebraska Press, 1961. Pp. 39-65.
- TEITELBAUM, P., & EPSTEIN, A. N. The lateral hypothalamic syndrome. *Psychological Review*, 1962, 69, 74-90.
- WALKER, A. E. *Posttraumatic epilepsy*. Springfield, Ill.: Charles C Thomas, 1949.
- WASMAN, M., & FLYNN, J. P. Directed attack elicited from hypothalamus. *Archives of Neurology*, 1962, 6, 220-227.
- WHEATLEY, M. D. The hypothalamus and affective behavior in cats: A study of the effects of experimental lesions, with anatomic correlations. *American Medical Association Archives of Neurology and Psychiatry*, 1944, 52, 296-316.
- WILLIAMS, D. R., & TEITELBAUM, P. Some observations on the starvation resulting from lateral hypothalamic lesions. *Journal of Comparative and Physiological Psychology*, 1959, 52, 458-465.

(Received January 29, 1964)

APPARENT REVERSAL (OSCILLATION) OF ROTARY MOTION IN DEPTH:

AN INVESTIGATION AND A GENERAL THEORY

R. H. DAY¹ AND R. P. POWER

University of Sydney

3 explanations of apparent reversals (oscillation) of rotary motion in depth attribute this effect to misjudgment of orientation. These explanations are based mainly on observations of a trapezoidal "window" in rotation. The experiments reported here show that perspective effects in a trapezoidal window do not increase reversal frequencies and that other shapes in addition to a trapezoid exhibit the effect with similar frequencies. The experiments also failed to confirm that misjudgments of orientation are a causal condition of apparent reversals. A general theory in terms of an identity of projected (retinal) motion characteristics for clockwise and anticlockwise motion is proposed with supporting evidence. Apparent orientation is held to be a consequence of rather than a necessary condition for apparent reversal. This theory is sufficiently general to explain apparent reversals ("fluctuations") in the orientation in depth of static figures and objects and to explain also the kinetic depth effect. All these phenomena are held to derive from an identity of retinal projections for 2 or more motions or orientations of an object in space.

Objects rotating in depth relative to an observer are frequently judged, usually with monocular vision, as apparently reversing their directions of rotary motion. Recession of a point or edge is sometimes reported as approach and vice versa. This effect, as far as is known, was first reported by Kenyon (1898) in the following terms:

A curious illusion connected with an ordinary two-winged pendant fan . . . consists in the fan appearing to rotate in the opposite direction. . . . Two other illusions . . . may be noted. In one the vanes, instead of rotating, seem to flap together; in the other the two arms appear to be continually withdrawing into and pushing out from the hanging rod [p. 371].

The effect was later mentioned by Miles (1929) and, more recently, extensively investigated by Ames (1951) who explained it in terms of assumptions consequent upon past experience

with certain shapes. Alternative explanations have since been proposed by Pastore (1952) and Graham (1963). All three explanations are based largely on observations of a rotating isosceles trapezoid² cut and painted to resemble a mullioned window in perspective. Quantitative data from a series of recent experiments (Day & Power, 1963), however, failed to support any of these explanations and a more general theory is presented here with supporting evidence.

It is intended to review briefly earlier explanations, to present experimental

² In an earlier paper (Day & Power, 1963) the term trapezium was used instead of trapezoid. This inconsistency is clarified by reference to the McGraw-Hill *Encyclopedia of Science and Technology*, Vol. 14, 1960: "Trapezoid: a term used in the United States for a quadrilateral with two sides parallel. In Great Britain such a figure is often called a trapezium, a term used in the United States for a general quadrilateral."

¹ Now at Monash University.

data which render them questionable, and then to describe the more general theory. It will also be shown that the present theory can be extended to explain a variety of apparent reversing phenomena associated with stationary and moving "ambiguous" figures of two and three dimensions.

CURRENT EXPLANATIONS OF APPARENT REVERSAL

Ames' Explanation

The explanation proposed by Ames (1951) derives from qualitative data and rests upon two basic assumptions. First, an isosceles trapezoid cut and painted to resemble a mullioned window in perspective is seen as slanted to the frontoparallel plane when, in fact, it lies in that plane. The longer vertical edge is seen as closer than the shorter irrespective of the degree of slant. Second, a cue for rotary motion in depth is held to be a continuous variation in the total horizontal angle subtended at the eye. Thus, when the window, actually in the frontoparallel plane, but seen as slanted to it, begins to rotate, an observer would expect the total horizontal angle to increase as the apparently further edge approaches. Since, however, the horizontal subtense *decreases*, and, since this is a cue for recession, the window's direction of motion is seen as opposite to its true direction. Ames (1951) argued that the windowlike properties of the shape enhanced the effect since a window is commonly rectangular. A trapezoidal shape, therefore, would tend to be seen as a slanted window. Presumably a door or book-cover pattern or some such would serve as well to emphasize the reversal effect.

Pastore's Explanation

Pastore (1952) demonstrated and Canestrari (1956) and Mulholland (1956, 1958) confirmed that reversals

of rotary motion were not confined to trapezoidal shapes. It was found by the former that the effect occurred with a variety of forms including a circle, an ellipse, a "lopsided" ellipse, a triangle, and a solid object. Pastore's explanation is based on a difference between true and apparent slant. When there is a discrepancy between these, apparent direction of motion is opposite to true direction. Contrariwise, when apparent slant conforms to true slant, perceived motion direction is the same as true direction. This principle was demonstrated by slanting a trapezoidal window at 45 degrees from the frontoparallel with the longer vertical edge furthest, obtaining a judgment of slant, and then rotating it through about 90 degrees to obtain a judgment of motion direction. Kilpatrick (1953) has pointed out that there is little to choose between the explanation proposed by Ames and that put forward by Pastore. Both stress the role of apparent slant in determining apparent reversals in rotary motion.

Graham's Explanation

Graham (1963) has explained apparent reversals in terms of an identity of differential angular velocities of points on the surface of the object for clockwise and anticlockwise movement, and the resolution of direction by perspective cues. It is argued that such ambiguity of movement parallax deriving from identical differential angular velocities is resolved by the use of perspective cues provided by the trapezoidal window. Thus, the short vertical edge, no matter in which quadrant of the circle it lies, is judged as being in one of the two far quadrants. When the short edge actually lies in one of these quadrants, movement direction is correctly judged, but when it lies in one of the two near quadrants, reversed rotation seems to occur. The similarity

between this explanation and that proposed by Pastore (1952) is recognized by its author.

TESTS OF CURRENT EXPLANATIONS

Although the three explanations reviewed above derive from different theoretical positions, they have two features in common. First, they are based almost entirely on data obtained from observations of a rotating trapezoidal shape. Pastore (1952) demonstrated reversals with a variety of shapes, but he argues from observations of a trapezoidal window slanted at 45 degrees from the frontoparallel. Second, all three explanations treat misjudgment of slant as a critical determinant of apparent reversal. That is, misjudgment of the frontal orientation of the object is regarded as a necessary condition for the oscillatory effect. The greater apparent distance of the short vertical edge of the trapezoid regardless of its true orientation, and the enhancement of this effect by window characteristics, is regarded as the basis of the apparent reversal phenomenon. In the following experiments, therefore, these stimulus properties have been examined.

Throughout what follows the term apparent reversal will be used to refer to the subject's report of the approach of an edge or part of the object when in fact it is receding and vice versa. The alternation of apparent reversals with reports of true motion direction is referred to as oscillation. The term (apparent) orientation will be used throughout to refer to the angle that the shape makes with the subject's frontoparallel plane when it is mounted on a vertical axis.

Experiment 1: Effects of Stimulus Shape and Pattern

In a series of recent experiments (Day & Power, 1963), six stimulus

conditions were derived from two shapes (isosceles trapezoid and rectangle) and three surface patterns (plain white, vertical black and white bars, and an Ames window). The vertical edges of the trapezoids were 11.75 and 6.25 inches, and their non-parallel edges 9.625 inches. The rectangular shapes were 11.625×9.25 inches. The trapezoidal window was mullioned and painted to resemble a window with thickness slanted to the line of regard. The objects, which were mounted in an enclosed chamber, rotated clockwise at 8 rpm and were viewed monocularly from 52 inches through a viewing tube. Six independent groups of 12 subjects drawn from undergraduate classes in psychology signaled apparent reversals in the direction of rotation by operating a press-switch during two 20-revolution trials. Reversals were recorded on an automatic counter.

Mean frequencies of reversals for each of the six shapes and patterns based on two trials are shown in Figure 1a. Although there were large and significant differences between reversal frequencies for the two shapes ($p < .001$), there was none between patterns ($p > .10$). The surface characteristics of a mullioned window in perspective failed to increase significantly the frequency of apparent reversals, a finding contrary to the generally held view that resemblance to a commonly rectangular object contributes to the effect. Identical results were obtained in a repetition of this experiment using subjects quite unfamiliar with the effect. In the case of the rectangular shapes only a small proportion of subjects in each group reported apparent reversals, whereas all subjects reported reversals with the trapezoidal shapes.

In a second experiment under the same viewing and responding conditions three shapes (circular, elliptical,

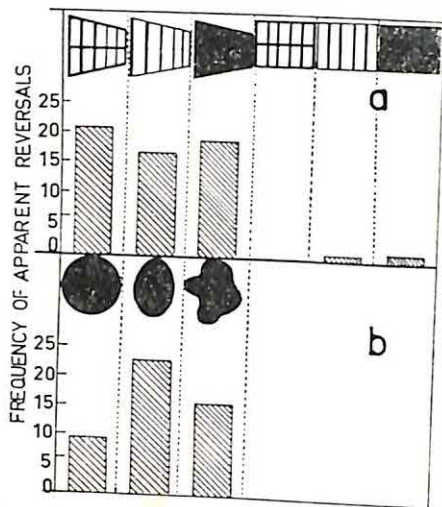


FIG. 1. Part a: Frequencies of apparent reversal of rotary motion for two shapes (trapezoidal and rectangular) and three surface patterns (plain white, vertical bars, and "window"). Part b: Frequencies of reversal of rotary motion for a circular, elliptical, and irregular shape. (In this and subsequent figures the plain white shapes are shown in black for convenience.)

and irregular) with a vertical length of 10 inches were observed during two 20-revolution clockwise trials by independent groups of 15 subjects. The histograms in Figure 1b show that whereas the elliptical and irregular shapes exhibited similar frequencies of reversal ($p > .05$), the circular shape did so less frequently. The apparent reversal frequency of the latter shape

differed significantly from that of both ellipse and irregular shape ($p < .01$).

Experiment 2: Apparent Orientation

Since there are no quantitative data on the apparent orientation of a trapezoidal window and other shapes when stationary, Experiment 2 (Power, 1964) was designed to obtain such data for two orientations of six shapes. Using the same monocular viewing conditions as in Experiment 1, six groups of 10 subjects matched the orientation of a trapezoidal window, an ellipse, a rectangular window, and three irregular shapes while they were stationary. One group was assigned to one shape and was required to match its orientation when it was slanted at 60 and 120 degrees from the sagittal (0 degree) plane. The shapes were suspended from the top of the viewing chamber, and the subject matched their orientation by means of a rectangular shape with a lattice pattern on its surface. The axes of the standard and comparison shapes were coextensive, and preliminary observations had shown that the comparison rectangle could be accurately oriented into the frontoparallel plane. Adjustment of the slant of the lower comparison rectangle was made by means of a hand

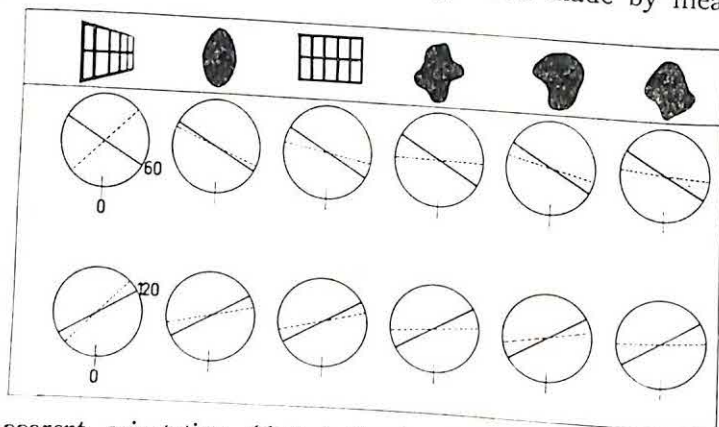


FIG. 2. Apparent orientation (dotted lines) of four shapes (trapezoidal window, rectangular window, ellipse, and irregular) at two angles (60 and 120 degrees) from sagittal (0 degree) plane.

wheel in front of the subject which was connected by means of a pulley and belt to the axis of the shape. The task assigned the subject involved adjusting the slant of the lower comparison shape so that it lay in the same apparent plane as the upper standard shape.

The mean angular matches for the six shapes at 60 and 120 degrees from the sagittal plane are shown in Figure 2 as dotted lines. The 60- and 120-degree true orientations are shown as heavy continuous lines, and the shapes themselves are shown accompanying each diagram. While the ellipse, the rectangular window, and the three irregular shapes were judged as oriented in the correct quadrants at both angles, the trapezoidal window was judged as being oriented with its shorter vertical edge in one of the two far quadrants regardless of its true orientation. Whereas this judgment was correct for the 60-degree position, it was incorrect for the 120-degree position.

Thus the trapezoidal window, whose mean apparent reversal frequency in Experiment 1 was 21, was misjudged as to its quadrants of orientation in one of its two positions. On the other hand, the orientation quadrants of the elliptical and irregular shapes whose mean reversal frequencies were 24 and 17, respectively, were correctly judged.³ These data suggest that misperception of orientation is not a necessary condition for the occurrence of apparent reversals of direction in rotary motion.

Experiment 3: Relationships between Apparent Orientation, Initial Movement Direction, and Reversals

The purpose of the third experiment (Power, 1964) was to investigate un-

der the same conditions the relationships between judgments of orientation, of initial direction of movement, and of reversals. The second of these (initial direction of movement) was included since Graham (1963) and Pastore (1952) argue that, in the case of the trapezoidal window, apparent orientation determines perceived direction of movement.

The apparatus and viewing conditions were similar to those of Experiment 2. The standard shapes were a trapezoidal window, an ellipse, an irregular shape, and a rectangular window all of which had been used in Experiment 1. For judgments of orientation the four shapes were placed in the subject's frontoparallel plane and the two asymmetrical objects (trapezoid and irregular) were positioned with one edge both left and right. Thus there were six conditions. Independent groups of 10 subjects drawn from introductory classes in psychology observed one shape during all three phases. In the first phase the subject adjusted the lower latticed rectangle so that it appeared in the same plane as the upper standard. Immediately this match was made the standard was rotated anticlockwise at 4.5 rpm and the subject indicated apparent direction of movement (Phase 2). In the third phase the subject indicated apparent reversals during a 20-revolution trial as in Experiment 1.

The results are shown in Table 1 for the four shapes and the two positions of the trapezoidal and irregular shapes.

Pastore's (1952) explanation states that, when the longer vertical edge of the trapezoid is judged nearer when in fact it is farther, rotation direction appears opposite to true direction. Judgments of initial motion direction immediately following orientation judgments (Table 1, fifth row) show that when the longer edge of the trapezoid

³ For the order in which they occur in Figure 2, the mean apparent reversal frequencies of the two irregular shapes which were not used in Experiment 1 were found to be 17 and 18 in a separate experiment.

TABLE 1
MEANS AND STANDARD DEVIATIONS OF ORIENTATION MATCHES AND REVERSAL
FREQUENCIES WITH NUMBER OF SUBJECTS JUDGING INITIAL
MOVEMENT AS CLOCKWISE (C) AND ANTICLOCKWISE (AC):
EXPERIMENT 3

		Shapes					
		Irregular (protuberance right)	Irregular (protuberance left)	Trapezoidal window (short side right)	Trapezoidal window (short side left)	Ellipse	Rectan- gular window
Orientation (Degrees from sagittal plane)	\bar{X}	94.6	91.1	320.9	38.3	103.5	92.4
	SD	5.06	5.63	10.3	12.66	32.35	3.67
Reversal frequency	\bar{X}	16.4	22.0	28.9	25.5	23.9	.2
	SD	9.1	12.65	9.33	7.17	8.07	.01
Initial movement direction	C	4	5	0	10	2	1
	AC	6	5	10	0	8	9

was judged nearer on the left, all judgments of direction were correct (i.e., anticlockwise). When, however, the longer edge was judged nearer on the right, all judgments of rotation direction were incorrect (clockwise). Thus, Pastore's (1952) statement is valid for the trapezoidal window. The phi coefficients between judged quadrants of orientation and judged initial direction of motion (clockwise or anticlockwise) have been computed for the four shapes, the two positions of the irregular and trapezoidal shapes being combined for this purpose. These coefficients are .297 for the irregular shape, .900 for the trapezoid, .234 for the ellipse, and .250 for the rectangle. The relationship between judged orientation and judgments of initial direction of motion achieves significance only in the case of the trapezoidal window and for all other shapes fails to reach significance. Thus, whereas this relationship holds for one shape (trapezoid) which manifests 29 and 25 apparent reversals during rotation, it does not do so for two other shapes which mani-

fest reversal frequencies of 16 and 22 (irregular) and 24 (ellipse).

Summary of Experimental Data

The data from the first experiment show that the apparent reversal frequency of a trapezoidal window is no different from the reversal frequencies of the same shape with alternative surface patterns. These data show also that, with the exception of rectangular and circular shapes, other shapes exhibit apparent reversals with about the same frequency as a trapezoid. The results from Experiment 2 demonstrate that, whereas the quadrants of orientation of a trapezoidal window are misjudged when the short vertical edge is near, the orientation quadrants of other reversing shapes are not. In the third experiment no evidence was found for a causal relationship between apparent orientation and apparent reversals, nor was there evidence for a relationship between apparent orientation and initial apparent direction of movement. Although the trapezoidal window is misjudged as to its orienta-

tion and also exhibits apparent reversals during rotation, the former is not necessary for the latter with the other shapes. In brief, the hypothesis that apparent orientation deriving from perspective cues is necessary for apparent reversals has not been sustained.

AN ALTERNATIVE EXPLANATION

In this and the following section a theory of apparent reversals of rotary motion in depth will be proposed. It will be further argued that the theory applies also to cases of reversal in static figures and objects.

When a plane shape rotates in depth relative to an observer, the retinal projection of its motion is a repeated expansion and contraction as the object moves from the frontoparallel to the sagittal plane of the observer. If the retina is considered as a plane surface then the motion of projected points follows a sine function. The retinal projection expands rapidly as the object moves from the sagittal plane ("edge-on" position) and slows, and finally ceases as it approaches and reaches the frontoparallel ("full-on" position).

Now, when the shape rotates clockwise from the sagittal to the frontoparallel plane, the retinal projection expands from its minimum width determined by the thickness of the object, to its maximum width determined by its horizontal extent. But when the shape moves anticlockwise between these planes an identical expansion of its retinal projection occurs. That is, there is an identity of retinal motion characteristics for clockwise and anticlockwise rotation. The same applies to the contraction phase as the object moves from the frontoparallel to the sagittal plane.

Graham (1963) has shown in a detailed analysis the identity of differential angular velocities for points on the

surface of a rotating plane during the approach and recession phases of rotary movement. It is clear from this treatment that the subject cannot judge whether these points are approaching in a near quadrant or receding in a far quadrant, since the same sign and quantity of differential angular velocity apply in each.

Apparent reversals in rotation direction can be explained in the following terms. Since there is an identity of retinal motion for the two directions of rotation, then, in the absence of cues to direction in depth such as retinal disparity, the object will occasionally be judged as moving clockwise and occasionally as rotating anticlockwise.

The data from these experiments indicate that the identity of retinal motion characteristics for approach and recession of the rotating shape is a necessary condition for the occurrence of apparent reversals. When this condition obtains, therefore, it would be expected that judgments of approach or of recession would be equally probable. That is, when the retinal projection of the shape changes from expansion to contraction (and vice versa), the subject will judge it as either approaching or receding, and there is no reason to suppose that other than chance factors determine this judgment. In support of this is the observation that the rotating shapes are not reported as reversing during every revolution. Reversals are reported along with complete revolutions and, as will be seen below (Experiment 4), about half the possible number of apparent reversals occur during a trial. Although the retinal identity of clockwise and anticlockwise motion is clearly recognized by Graham (1963), he nevertheless attributes the apparent reversing effect to apparent orientation determined by cues to perspective. It is pertinent, however, to raise the issue concerning this rela-

tionship between apparent orientation in depth and apparent reversals. If the object appears to rotate, say, anti-clockwise, then it would be expected that apparent orientation at any one moment would be in accord with that direction and vice versa. That is, apparent orientation can be regarded as a consequence rather than as a cause of apparent direction of movement. The data from Experiments 2 and 3 are convincing in showing that apparent orientation does not determine either reversals or initial direction of movement. This relationship is further discussed below.

The principle involved in this explanation is simply that an identity of retinal motion for two directions of rotation leads to an ambiguity resulting in two possible judgments of direction. It would be expected, however, that if cues to motion direction were available, then apparent reversals would either be reduced in frequency or eliminated. One such cue is that provided by retinal disparity which gives information about the true orientation of the object in depth from moment to moment. Apparent reversals are not reported with binocular viewing unless the viewing distance is greater than that within which retinal disparity is normally operative (Ames, 1951).

Further cues to motion direction are probably provided by certain shapes which exhibit few or no reversals. Although the role of these cues has yet to be confirmed, it is reasonable to discuss their possible modes of operation here. If, as with rectangles, at some stage during rotation the vertical edges project equal visual angles, then decrease of one angle with recession, and increase of the other with approach, could provide directional cues. That is, equality of retinal projections when frontoparallel provides information for

direction discrimination; information not available when the two vertical edges are unequal so that at no stage of rotation are "base-line" conditions for comparison available. Support for the role of this cue is provided by Mulholland's (1956) data showing that apparent reversals occur predominantly with asymmetrical objects. It is not possible, however, to speculate with conviction on the directional cues responsible for the reduced frequency of reversal of the circular shape (Experiment 1). Although this shape reversed with a relatively high frequency, it did so less than the trapezoidal, elliptical, and irregular shapes.

ANGLES OF APPARENT REVERSAL

The theory of apparent reversal which has been set out here states that the effect derives from an identity of horizontal retinal motion projections for the approach and recession of an object rotating in depth. It follows that if an edge or point is judged as receding when in fact it is approaching, then the orientation of the object will also be misjudged. That is, the apparent orientation of the object will be determined by the apparent direction of motion. Thus, the theory is in this regard opposite to those of Ames (1951), Graham (1963), and Pastore (1952) insofar as apparent orientation is held to be a consequence of, rather than a necessary condition for, apparent reversals. In brief, it is contended that whichever direction of rotation is judged, both being equally probable, then orientation will be judged accordingly.

It also follows that apparent reversals would be expected to occur when the object is nearly frontoparallel or nearly sagittal. As the object passes through the frontoparallel the retinal projection commences to decrease in horizontal extent, and as it passes

through the sagittal plane retinal expansion occurs. At these points of change an observer can make one of two judgments of rotation direction. At all other orientations reversals would not be expected since these would necessarily result in abrupt changes in apparent orientation. For example, if the right edge of a shape is judged as approaching when about 45 degrees from the frontoparallel, an apparent reversal would require a sudden jump of the edge to 135 degrees (the opposite orientation). If a reversal were to occur at this point without such an abrupt reorientation, then the object would necessarily appear to expand as the edge receded. There is no evidence for such sudden changes in apparent orientation. Thus, as the object passes through the frontoparallel and sagittal planes of the observer, changing phase from expansion to contraction, apparent direction of rotation may be judged as clockwise or anticlockwise.

Experiment 4: Angles of Orientation at Which Apparent Reversals Occur

The fourth experiment was conducted to test the prediction that apparent reversals occur when the object is in either the sagittal or frontoparallel planes of the observer.

Using the same apparatus and viewing conditions as in Experiment 1, the subjects were required to press the switch when apparent reversals occurred. These responses were recorded on a constant speed paper recorder. Also recorded were the occasions when one edge of the shape was near and exactly sagittal (0 degree). The distance between these latter signals represented 360 degrees, and it was therefore possible to establish the angular positions of the shapes at which apparent reversals occurred.

Three groups of five subjects each

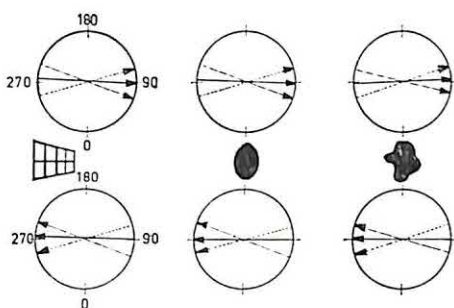


FIG. 3. Mean angles of apparent reversal for three shapes (trapezoidal window, ellipse, and irregular) rotating clockwise (dotted and dashed lines) and anticlockwise (dotted lines). (The mean reversal angles for the two directions of rotation combined are also shown, continuous lines. The apparent reversal angles are shown separately for those occurring near 90 degrees, upper row, and near 270 degrees, lower row.)

underwent 10 10-revolution trials, 5 in which rotation was clockwise and the other 5 in which it was anticlockwise. One group observed the ellipse, another the trapezoidal window, and the third the irregular shape. All shapes were those used in Experiment 1, and the viewing and responding conditions were also the same.

The mean angles of reversal are shown in Figure 3 for clockwise and anticlockwise rotation and for the two directions taken together. It can be observed that the mean reversal angles for the two directions combined approximate closely to the true frontoparallel (90 and 270 degrees), while those for the two directions considered singly occur beyond and before the frontoparallel. Of the 15 subjects, there were only 2 who signaled reversals near the sagittal position. These are not shown in Figure 3.

In view of the ambiguity of the stimulus conditions and the minimum velocity of the retinal projection when the object is frontoparallel, it is reasonable to attribute the difference between apparent reversal angle and frontoparallel for the clockwise and counterclockwise

THE PRINCIPLE OF BELIEF CONGRUENCE AND THE CONGRUITY PRINCIPLE AS MODELS OF COGNITIVE INTERACTION¹

MILTON ROKEACH

AND

GILBERT ROTHMAN

Michigan State University

University of California, Berkeley

The principle of belief congruence, previously presented in Rokeach (1960), is elaborated here so that it can be applied to instances of cognitive interaction. The belief congruence model is compared and contrasted with Osgood and Tannenbaum's congruity principle, the main difference between the 2 principles being that the former asserts that the outcome of cognitive interaction cannot be accurately predicted from a knowledge of the evaluative meaning of the 2 elements judged in isolation, while the latter asserts that it can. Quantitative predictions generated by the 2 models are compared with obtained results. It is found that the average absolute error of the congruity model is 1.07 and the average absolute error for the belief congruence model is .34, thus suggesting that the latter model's predictions are approximately 3 times as accurate as the congruity model's predictions.

The main purpose of this paper is to elaborate further the principle of belief congruence (Rokeach, 1960) so that it will apply to instances of cognitive interaction. A second purpose is to compare and contrast the predictions generated by this principle with those generated by Osgood and Tannenbaum's (1955) congruity principle. It will be suggested that the two principles cannot be equally valid since they appear to lead in many instances to contradictory predictions. A third purpose is to present the results of empirical work designed to determine which set of predictions is the more accurate, and a fourth purpose is to try to reconcile the data and interpre-

tations presented by Osgood and his co-workers with those in our own research program.

THE PRINCIPLE OF BELIEF CONGRUENCE

The principle of belief congruence asserts that we tend to value a given belief, subsystem, or system of beliefs in proportion to their degree of congruence with our own belief system and, further, that we tend to value people in proportion to the degree to which they exhibit beliefs, subsystems, or systems of belief congruent with our own. Congruence can be defined both in terms of similarity and importance. Given two beliefs or subsystems of belief equal in importance, the one more similar to our own is the more congruent; conversely, given two beliefs or subsystems perceived to be equally similar to our own, the one judged as more important is the more congruent with our own belief system.

Before discussing how the principle of belief congruence applies to the area of cognitive interaction, it is first neces-

¹ The research described herein was supported by a grant from the National Science Foundation, and an early draft of this paper was prepared by the senior author at the Center for Advanced Study in the Behavioral Sciences. The final version of this paper was written following a stimulating exchange of views with Charles E. Osgood. We are grateful to Osgood for his criticisms, but this acknowledgement should not be taken to mean that he necessarily concurs with the formulations and conclusions stated herein.

sary to assume that any stimulus (a verbal concept, percept, or event) has the property of activating within a person only a portion of his belief system—that portion which is relevant to or associated with the stimulus. Which particular portion and how broad a portion of the belief system is activated depends on the particular stimulus presented. A particular stimulus may activate only a single belief or a whole subsystem of beliefs varying in breadth or in the number of inter-related component beliefs. KHRUSHCHEV or COMMUNISM, for instance, may activate a broader area of the person's belief system than IVAN PETROV or COLLECTIVE FARM.

Moreover, a person should typically have no difficulty, when presented with an isolated stimulus, in assessing the direction and importance of the beliefs it activates along any given scale (say, a semantic differential scale) because his previously learned belief system should provide him with a generalized frame of reference for judging direction and importance.

Consider now what may reasonably be expected to happen when one stimulus is associated with another through some assertion. The linkage is assumed to give rise within the person to a unique configuration activating comparison processes designed to determine whether or not there is to be cognitive interaction and, if so, its outcome. An assertion, following Osgood, may be associative or dissociative and may take one of four forms: simple linguistic qualification (LAZY ATHLETE), simple perceptual contiguity (an advertisement showing a MOVIE STAR smoking CAMELS), statements of classification (JONES is a PHYSICIST), and source-object assertions (DE GAULLE opposes TEST BAN). Cognitive interaction refers to the process by means of which a single evaluative

meaning emerges as a result of combining two stimuli, each having their separate meanings.

All the four types of assertions mentioned above have something in common: they are unique configurations cognitively representing a *characterized subject* (*CS*)—a person, thing, or idea characterized or qualified in some unique way. The unique configuration consists of two components: a *subject* (*S*), capable of being characterized in many ways, and a *characterization* (*C*), capable of being applied to many subjects.

Let us now restate the principle of belief congruence in order to apply it to all conditions which might and might not lead to cognitive interaction and at the same time to quantitatively predict its outcome. Whenever two stimuli are brought into association with one another through an assertion they form a unique configuration activating two kinds of comparison processes: the stimuli will first be compared for mutual relevance, and if they are perceived to be at least partially relevant for one another, they will then be compared for relative importance.

Comparison of Relevance

Upon presentation of *CS* a person will first ascertain whether or not the two components, *C* and *S*, are relevant for one another. If he judges them to be not relevant, there are no psychological grounds for expecting cognitive interaction. The irrelevant component should be ignored; that is, it should exert no influence on the evaluation of *CS*. Consequently, *CS* should be evaluated in about the same way the remaining component is evaluated.

Comparison of Relative Importance

More interesting for psychological theory and research are those instances wherein a person judges *C* and *S* to

be at least partially relevant for one another.

Relative Importance of C and S. Under the condition of at least partial relevance, a person will next compare the two components for relative importance with respect to another, judged within the general frame of reference provided by one's previously learned belief system. And, we conjecture, the evaluation of the configuration *CS* should be a simple average of the evaluations of *C* and *S* considered separately, weighted by the perceived importance of *C* and *S* relative to one another within the context *CS*:

$$d_{CS} = (w)d_C + (1 - w)d_S \quad [1]$$

where d_{CS} , d_C , and d_S refer, respectively, to the degree of polarization (positive or negative) of the characterized subject, the characterization, and the subject, and where (w) and $(1 - w)$ refer to the perceived importance of d_C and d_S relative to one another in the context *CS*.

Relative Importance of CS and C. When *C* reaches 100% in importance (and *S* 0%) we must posit the activation of yet an additional comparison process over and above a comparison of *C* and *S*, namely, a further comparison of the relative importance of *CS* and *C*. Suppose a person perceives IRRESPONSIBLE as all-important in the context IRRESPONSIBLE FATHER. In such a case, it might be reasonable at first glance to assume from Formula 1 that a person's evaluation of *CS* would be completely determined by his evaluation of *C*; that is, he evaluates *CS* the same way he evaluates *C*. Such an assumption, however, overlooks the possibility that by virtue of the interaction between *C* and *S* within the framework of one's total belief system a person's evaluation of *CS* may be even more extreme than his evaluation of *C*. He may feel strongly negative

toward IRRESPONSIBLE but even more strongly negative toward IRRESPONSIBLE FATHER because he may feel that fathers especially ought not to be irresponsible. In other words, he may evaluate *CS* as falling outside the limits of *C* and *S*. A characteristic which is negatively evaluated in isolation may be judged even more negatively when it is lodged in a positive subject; conversely, a characteristic which is positively evaluated in isolation may be judged even more positively when it is lodged in a negative subject.²

The combined effects of the two comparison processes (*C* versus *S*, and *CS* versus *C*) on the evaluation of *CS* under the condition described may be quantitatively expressed as follows:

$$d_{CS} = d_C + (v)d_C \quad [2]$$

where d_C represents the effect of the first comparison process (*C* versus *S*) and where $(v)d_C$ represents the additional effect of the second comparison process (*CS* versus *C*), and where v represents the extent to which the per-

² We are dealing here with a comparison process sometimes referred to as *overassimilation*, a process whereby a stimulus not only takes on the valence of another stimulus with which it is associated, but in addition takes on an even stronger valence. This phenomenon can often be observed in everyday life: (a) A Jew who converts to Christianity is sometimes regarded as being "worse" than a Goy; (b) a positively evaluated person who defects to the other side is even more severely condemned than a negatively evaluated person on the other side (attitude toward the renegade, Rokeach, 1960); (c) a negatively evaluated person who defects to our side is even more warmly embraced than those already on our side.

All the above instances involve a comparison between *CS* and *C*. For the sake of completeness the converse possibility should also be mentioned, namely, a comparison process involving the relative importance of *CS* and *S* when *S* is perceived to be 100% important. On intuitive grounds, however, this does not appear to us to be psychologically meaningful and will therefore not be given further consideration here.

son attaches greater importance to CS than he does to C . When the person judges CS to equal C in importance, $v = 0$ and $d_{cs} = d_c$. When the person judges CS to exceed C in importance, v will equal some coefficient expressing the extent to which CS is perceived to exceed C in importance, and d_{cs} will exceed d_c by the amount $(v)d_c$. When d_c is positive, d_{cs} will be more positive, and when d_c is negative d_{cs} will be more negative.

One important restriction should be noted here. The value of d_{cs} cannot be allowed to exceed the most extreme score on whatever happens to be the scale of measurement employed in empirical research (e.g., ± 3 on a scale ranging from -3 to $+3$). Consequently, $d_c + (v)d_c$ cannot be allowed to exceed this extreme score either. If d_c , for example, already equals the most extreme score then the sum of d_c and $(v)d_c$ is arbitrarily assigned the same score.

CONTRAST BETWEEN THE BELIEF CONGRUENCE AND CONGRUITY PRINCIPLES

The formulation of the principle of belief congruence differs from Osgood and Tannenbaum's congruity principle in several important respects:³

1. The congruity principle is an additive model which predicts the outcome of cognitive interaction solely from a knowledge of the direction and degree of polarization of the two stimuli considered in isolation. In contrast, the principle of belief congruence is a configurationist model asserting that the comparison processes cannot become activated until the two stimuli are linked together to form a unique

³ The most detailed exposition of the congruity principle will be found in the original article by Osgood and Tannenbaum (1955). For further elaborations, see Osgood, Suci, and Tannenbaum (1957) and Osgood (1960). An excellent summary of this work is given by Brown (1962).

gestalt (Asch, 1946, 1952) and, consequently, the outcome of the cognitive interaction cannot be accurately predicted solely from a knowledge of the direction and intensity of the two stimuli considered separately.

This is not to say, however, that the evaluative meaning of the unique configuration, CS , is unpredictable. As already noted, CS is evaluated against the general background of one's total belief system. Elsewhere, Rokeach (1960, 1964, 1965) has attempted to describe various types of beliefs within the belief system and to locate them along a central-peripheral dimension. These formulations provide us with a basis for predicting, at least roughly, the evaluative meaning of any given CS : they guide educated guesses about the relative location of beliefs activated by C versus S , and by CS versus C , along a central-peripheral dimension.

2. The two models differ in their conception about the psychological meaning of incongruity. Exactly what is incongruous with what? According to Osgood and Tannenbaum (1955) the incongruity is between the C and the S , and the greater the disparity between C and S the greater the incongruity and the greater the pressure to reduce it when the assertion is positive; the reverse holds when the assertion is negative. Reduction of incongruity between C and S is achieved by a compromise in which C and S both change toward or away from one another in inverse proportion to their respective degrees of polarization.⁴

⁴ The formula Osgood, Suci, and Tannenbaum (1957) use to predict the outcome of congruence interaction is:

$$d_{cs} = \frac{|d_c|}{|d_c| + |d_s|} (d_c) + \frac{|d_s|}{|d_c| + |d_s|} (d_s)$$

"where $|d|$ is deviation or polarization from neutrality on the scales regardless of sign, d is deviation from neutrality with respect to sign [p. 278]."

The psychological meaning of incongruity is formulated somewhat differently within the belief-congruence framework. The incongruity arises not from the psychological disparity between *C* and *S* but from the disparity between *C* and *CS*, or between *S* and *CS*, or both. The evaluative meaning of *CS* must somehow be made maximally congruent with one's belief system, which includes within it previously learned evaluative meanings of *C* and *S*. Since *CS* cannot be completely congruent with both *C* and *S* (assuming some discrepancy in the evaluative meaning of *C* and *S*), the question arises about the exact process whereby the evaluative meaning of *CS* will become maximally congruent with *C* and *S*. The principle of belief congruence attempts to describe how this outcome comes about. The more discrepant the relative importance of *C* and *S* with respect to one another the greater the pressure to evaluate *CS*, positively or negatively, like *C* or like *S*, whichever is the more important, but not both. The more *C* equals *S* in importance, the more equalized the pressure from *C* and *S*.

3. The congruity principle is essentially a compromise model, the only exception being the following: when an extremely polarized stimulus is positively associated with a neutral stimulus the meaning of the combined stimuli is assimilated to that of the extremely polarized stimulus. In contrast, the principle of belief congruence allows for various degrees of compromise, and assimilation and over-assimilation depending on the relative importance of *C* versus *S* and of *C* versus *CS*, in the context *CS*, regardless of the degree of polarization of *C* and *S*, considered separately.

4. The congruity model formally posits an assertion constant in the case of source-object assertions: there is a

greater force acting on the object than on the subject; that is, according to Brown (1962), "the object of the bond will be *more* affected than the source of the bond [p. 26]." In contrast, the principle of belief congruence denies the necessity of such an assertion constant and states instead that the magnitude of the force acting on the *source* and *object* depends on the relative importance of the beliefs activated by source and object. For example, in the assertions KENNEDY *praises* GOD and GOD *praises* KENNEDY, the congruity principle's assertion constant should lead to the expectation that there will be a greater average change of value on the two objects than on the two sources. In line with the principle of belief congruence we would expect instead that the less important component, whether source or object, would be the more affected.

A TEST OF THE TWO PRINCIPLES

Before presenting the results of empirical work specifically designed to test the contradictory predictions generated by the two principles, attention should first be drawn to a relevant and substantial body of data already made available by Rokeach, Smith, and Evans (1960), by Byrne and Wong (1962), and by Stein, Hardyck, and Smith (1965). The first study cited clearly shows that white subjects in the North and the South rate incongruous race-belief configurations such as *Negro who believes in God* and *white person who is an atheist* not by compromise, as the congruity model would predict, but by assimilation, the evaluation of the total configuration *CS* being more or less completely assimilated to that of *C*—the belief said to characterize the Negro or white. Similar results were obtained for Jewish children when they were presented with congruous and incongruous configura-

tions depicting Jews and Gentiles holding various beliefs. The Jewish children positively evaluate Gentiles who agree with their views (e.g., Gentile who is for Israel) and negatively evaluate Jews who disagree with their views (e.g., Jew who is against Israel). The total configurations (*CS*) are evaluated in essentially the same way the belief is evaluated in isolation (*C*) and, furthermore, regardless of whether the person holding the belief is a Jew or Gentile (*S*). Similarly, Byrne and Wong (1962) and Stein, Hardyck, and Smith (1965) presented fictitious personality profiles of whites and Negroes to their white subjects, some profiles being similar to and some different from the subjects' own profiles. The subjects evaluated positively those characterized as having personality profiles similar to their own and evaluated negatively those characterized as having profiles different from their own, regardless of whether the profile belonged to a white or Negro.

We may tentatively conclude from all the preceding that the subjects are responding in accord with the principle of belief congruence, which predicts assimilation whenever extremely important characteristics (beliefs or traits) are associated with positively and negatively valued people and not in accord with the congruity principle, which predicts compromise. But these studies do not provide us with a definitive test of the two principles because they employed research designs somewhat different from those of Osgood and his co-workers, and they did not employ the semantic differential. We therefore designed a new study involving the semantic differential which replicated Osgood and Ferguson's word combination study (Osgood et al., 1957, p. 275) except in one significant respect: we deliberately tried to select various concepts which, when presented

separately, would activate more or less highly polarized beliefs, thus ensuring that the predictions regarding the outcome of cognitive interaction by the principle of belief congruence would be diametrically opposed to those made by the congruity principle. As in the Osgood and Ferguson study, we selected components whose combinations were assumed to be credulous, and the purpose of the study was the same as their's, namely, the prediction of the evaluative meaning of word combinations.⁵

Forty-two white subjects, enrolled in an introductory psychology course at Michigan State University in the summer of 1961, rated 22 component concepts and 12 combinations shown in Tables 1 and 2 on three semantic differential scales representing evaluation (*valuable-worthless*; *admirable-deplorable*; *good-bad*). We used three types of assertions: simple linguistic qualification, statements of classification, and source-object assertions. The subjects first rated the 22 individual concepts on each of the three scales, and then rated 12 assertions linking 2 concepts. For the exact procedure see Osgood, Suci, and Tannenbaum (1957, p. 275). Predicted evaluation scores ranging from 1 to 7, which is, of course, equivalent to a range from -3 to +3, were calculated for each subject separately on the basis of the congruity formula (see Footnote 4) and, then, the mean predicted (from the congruity model) and obtained scale scores were calculated for all subjects.

Table 1 shows that the obtained means are reasonably close to the congruity model's predicted means in only 4 out of the 12 combinations (NEGRO

⁵ Actually, we obtained data on potency and activity as well as on evaluation. Since the evaluative dimension is of major concern to attitude theory and research, only the evaluation data will be reported here.

TABLE 1

MEAN ERROR OF PREDICTIONS AND SIGNIFICANCE TESTS OF DIFFERENCES BETWEEN MEAN OBTAINED EVALUATION SCORES FOR COMBINED CONCEPT AND MEANS PREDICTED BY THE CONGRUITY MODEL AND THE BELIEF CONGRUENCE MODEL

Assertion		M	SD	Mean error	t
1. A WHITE PERSON who is a COMMUNIST	Predicted: congruity	3.75	1.41		
	Obtained	2.06	1.05	1.69	8.09 ^{a***}
	Predicted: belief congruence	2.50	.65	.44	2.44 ^{b*}
2. A WHITE PERSON who is an ATHEIST	Predicted: congruity	4.43	1.43		
	Obtained	2.92	1.46	1.51	7.40 ^{***}
	Predicted: belief congruence	3.43	.43	.51	2.22 [*]
3. A NEGRO who believes in GOD	Predicted: congruity	6.32	.97		
	Obtained	6.42	.85	.10	.83
	Predicted: belief congruence	6.35	.28	.07	.51
4. A NEGRO who is an ANTICOMMUNIST	Predicted: congruity	5.96	.98		
	Obtained	5.89	1.24	.07	.50
	Predicted: belief congruence	5.91	.12	.02	.10
5. MY MOTHER IS INSINCERE	Predicted: congruity	4.42	.77		
	Obtained	2.04	.96	2.38	13.84 ^{***}
	Predicted: belief congruence	2.46	1.60	.42	1.62
6. UNIVERSITY PROFESSOR favors EXTRAMARITAL SEXUAL RELATIONS	Predicted: congruity	4.37	1.40		
	Obtained	2.90	1.56	1.47	5.90 ^{***}
	Predicted: belief congruence	2.89	.75	.01	.04
7. CLARK GABLE was in favor of FIDEL CASTRO	Predicted: congruity	3.56	1.21		
	Obtained	2.62	1.23	.94	4.14 ^{***}
	Predicted: belief congruence	2.73	.84	.11	.50
8. DISHONEST ATHLETE	Predicted: congruity	3.36	.88		
	Obtained	1.80	.85	1.56	9.28 ^{***}
	Predicted: belief congruence	2.06	.85	.26	1.49

TABLE 1—Continued

Assertion		M	SD	Mean error	t
9. UNFAITHFUL	Predicted: congruity	3.73	.77		
				1.60	8.54***
ROMANCE	Obtained	2.13	1.19		
				.39	1.63
	Predicted: belief congruence	2.52	1.11		
10. NIKITA KHRUSHCHEV	Predicted: congruity	5.25	1.09		
				.17	.74
advocates CLOSE	Obtained	5.08	1.57		
				.46	1.64
FAMILY TIES	Predicted: belief congruence	5.54	1.12		
11. RUSSIA extends	Predicted: congruity	5.66	1.10		
				.24	.80
FREEDOM OF	Obtained	5.42	1.70		
				.13	.45
THE PRESS	Predicted: belief congruence	5.55	.88		
12. A PROSTITUTE	Predicted: congruity	3.90	1.35		
				1.05	5.97***
who looks like	Obtained	2.85	1.58		
				1.25	3.89***
GRACE KELLY	Predicted: belief congruence	4.10	1.53		

* t tests for correlated measures between obtained and predicted (congruity) scores.

^b All t's between obtained means and means predicted by the belief congruence model were corrected for heterogeneity of variance, whenever necessary (Edwards, 1960, pp. 106-108).

* $p < .05$.

*** $p < .001$.

who believes in GOD; NEGRO who is an ANTICOMMUNIST; NIKITA KHRUSHCHEV advocates CLOSE FAMILY TIES; RUSSIA extends FREEDOM OF THE PRESS). Statistical tests, shown in Table 1, indicate that the obtained and predicted means for these four assertions are not significantly different from one another, results required by the congruity model. On the other eight assertions, however, the obtained means deviate markedly from the congruity model's predicted means, and in all eight instances are significantly different from one another at the .001 level.

It will be further noted in Table 2 that in the majority of the assertions the evaluative meaning of *CS* deviates markedly from that of *S* and adheres closely to that of *C*. In 6 out of the

12 assertions (Assertions 2, 3, 4, 5, 6, 7) *CS* does not differ significantly from *C*. In one instance (*WHITE PERSON who is a COMMUNIST*), *CS* is even more negatively evaluated than *C*, and significantly so—a clear instance of overassimilation. Finally, in Assertions 8 and 9, the evaluative meaning of *CS* is almost but not quite completely assimilated to *C*. When we consider the results of Assertions 1 through 9 all together, it is seen that the obtained results for *CS* are generally poorly predicted by the congruity principle.⁶ It is only in the

⁶ In only two of these nine assertions—the two Negro configurations—do the obtained means differ insignificantly from the congruity model's predictions (Table 1). The closeness of predictions in these two instances may be spurious; the disparity between *S* and *C* is very small and, consequently, it is

TABLE 2

MEAN DIFFERENCES BETWEEN THE EVALUATION OF THE COMBINED
CONFIGURATION AND EACH OF THE TWO COMPONENTS

	<i>M</i>	Mean difference	<i>t</i>
1. WHITE PERSON	5.35		
A WHITE PERSON who is a COMMUNIST	2.06	3.29	14.95***
COMMUNIST	2.45	.39	2.60*
2. WHITE PERSON	5.35		
A WHITE PERSON who is an ATHEIST	2.92	2.42	8.96***
ATHEIST	3.24	.32	1.78
3. NEGRO	5.38		
A NEGRO who believes in GOD	6.42	1.04	6.50***
GOD	6.52	.10	.91
4. NEGRO	5.38		
A NEGRO who is an ANTICOMMUNIST	5.89	.51	2.83**
ANTICOMMUNIST	5.96	.07	.39
5. MY MOTHER	6.51		
MY MOTHER is INSINCERE	2.04	4.47	19.60***
INSINCERE	1.86	.18	1.38
6. UNIVERSITY PROFESSOR	5.85		
UNIVERSITY PROFESSOR favors EXTRA- MARITAL SEXUAL RELATIONS	2.90	2.95	8.94***
EXTRAMARITAL SEXUAL RELATIONS	2.65	.25	1.32
7. CLARK GABLE	5.26		
CLARK GABLE was in favor of FIDEL CASTRO	2.62	2.64	9.78***
FIDEL CASTRO	2.36	.26	1.37
8. ATHLETE	5.84		
DISHONEST ATHLETE	1.80	4.04	20.20***
DISHONEST	1.53	.27	2.25*

TABLE 2—(Continued)

	<i>M</i>	Mean difference	<i>t</i>
9. ROMANCE	6.09		
UNFAITHFUL ROMANCE	2.13	3.96	17.22***
UNFAITHFUL	1.64	.49	2.97**
10. NIKITA KHRUSHCHEV	3.19		
KHRUSHCHEV advocates CLOSE FAMILY TIES	5.08	1.89	7.56***
CLOSE FAMILY TIES	6.43	1.35	5.40***
11. RUSSIA	4.08		
RUSSIA extends FREEDOM OF THE PRESS	5.42	1.34	4.19***
FREEDOM OF THE PRESS	6.15	.73	2.52*
12. PROSTITUTE	2.33		
PROSTITUTE who looks like GRACE KELLY	2.85	.52	3.25**
GRACE KELLY	5.67	2.82	10.68***

* $p < .05$.
 ** $p < .01$.
 *** $p < .001$.

case of Assertions 10 and 11 that the obtained evaluations are clearly predicted by the congruity model (Table 1). Finally, it is noted that only in the case of Assertion 12 is the mean evaluation of *CS* (PROSTITUTE *who looks like* GRACE KELLY) closer to *S* than to *C*, and here too the obtained mean differs significantly from the congruity model's prediction.

Let us now consider the extent to which the obtained results of *CS* are predicted by the principle of belief congruence which asserts that it is the relative importance of the two components perceived with respect to one another which is the crucial determi-

virtually impossible for the obtained results, assuming that they fall somewhere in between the component means, to differ significantly from the results predicted by the congruity principle.

nant of the evaluative meaning of the combined configuration. To generate specific predictions from this theory it is necessary to have independent measures of relative importance. Since we were not able to obtain such measures on the original sample of 42 subjects (for the simple reason that at the time of this study we had not yet fully developed the model presented here) we obtained the required measures of relative importance on a second comparable sample of introductory psychology students tested in the fall of 1963.⁷ We proceeded on the assumption, (a) that the average evaluative meaning of the components considered separately (shown in Table 2) would be approximately the same in the two samples,

⁷ We wish to thank Jacob Jacoby for his invaluable help in collecting and analyzing these data.

importance of *C* and *S* in percentage terms. However, if the subject responded "Yes" to Question 2a he proceeded to Question 2b, thus permitting a determination of whether *CS* was more important than *C*, and if so, how much more important.

The total number of subjects was 71. Of these 14 were eliminated because they obviously did not understand the instructions: they left one or more questions unanswered, or the two percentage estimates given in Question 2c did not add up to 100%, or they proceeded to answer Question 2b when they should have answered Question 2c, or vice versa.

Since we did not obtain semantic differential scores for the components for these subjects we assumed each subject as having the mean component scores shown in Table 2. We then computed a predicted score for each combination (*CS*) for each subject following Formulas 1 or 2, whichever was appropriate. It is the distribution of these predicted scores for *CS*, derived from one group, which were then compared with the obtained *CS* scores, derived from a comparable group.

It is necessary to emphasize that the results shown in Table 1 are group means. Predictions were actually made for each subject and for each assertion separately. In 38% of all the assertions (a total of 684 assertions) the subjects judged *C* to be 100% important (and *S* 0% important) in determining the meaning of *CS* (assimilation); in 16% of all assertions, the subjects further judged *CS* to be more important than *C* (overassimilation) in varying degrees.

It will be seen that the obtained means conform on the whole reasonably closely to those predicted by the belief-congruence model. The obtained means differ significantly from the predicted means in only 3 of the 12 asser-

tions (1, 2, and 12), and on the remaining 9 the obtained means do not differ significantly from the predicted means. In contrast, only 4 of the 12 obtained means differ insignificantly from the congruity model's predicted means.

A more sensitive test of the predictive power of the two models is obtained by comparing the mean absolute error of the two sets of predictions—the difference between obtained and predicted means. On Assertion 1, for example, the congruity model's mean error is 1.69 while the belief-congruence model's mean absolute error is .44. The mean error of prediction for all 12 assertions is 1.07 for the congruity model but only .34 for the belief-congruence model. Thus, the belief-congruence model's average error is only about one-third (.34) that of the congruity model's (1.07).⁹

One possible objection to our procedure is that the predicted and obtained *CS* means were not derived from the same subjects. We therefore also compared the predicted means with the means obtained from Question 1 (see instructions) in which the same subjects evaluated *CS* on a simple 7-point rating scale ranging from "strongly disapprove" to "strongly approve." In other words, obtained means were derived from the subjects' responses to Question 1, and predicted means were derived from their responses to Question 2.

The overall results are highly similar to those already presented (Table 1). The mean error of prediction for

⁹ On the first preliminary study designed to measure relative importance, the mean error for all 12 assertions was .56 ($N = 11$), and on the second preliminary study the mean error was .54 ($N = 63$). The final study, on which the mean error was .34, differs from the preliminary studies in one major respect: it included an evaluative rating of *CS* (see Question 1 on instructions) as well as ratings of relative importance.

Assertions 1 through 12 are, respectively: .34, .27, .04, .06, .77, .01, .55, .10, .07, .63, .53, and .82. For all 12 assertions the mean error of prediction is .35.

FURTHER CONSIDERATIONS CONCERNING THE VALIDITY OF THE TWO PRINCIPLES

It is clear that the data presented here are considerably more in line with the belief-congruence principle than with the congruity principle. Moreover, the relatively large size of the congruity model's prediction errors would appear to be at variance with the results previously presented by Osgood and Tannenbaum (1955) and by Osgood, Suci, and Tannenbaum (1957) in support of their claims regarding the congruity principle's general validity. In this section of the paper we will try to reconcile the present findings and interpretations (with respect to the two principles) with those put forward by Osgood and his co-workers, by looking more closely at the various kinds of evidence they have presented in support of the congruity principle. It will be suggested that these previously published data provide no more support for the congruity principle than do the present data.

1. Osgood, Suci, and Tannenbaum (1957) write concerning their word combination study,

One estimate—perhaps the crudest—is how often the obtained factor scores for the combinations fall between the factor scores for the components, a result required by the congruity formula [p. 280].

In their attitude change study Osgood and Tannebaum (1955) determine how often

predicted positive changes (+) and predicted negative changes (−) show corresponding signs in the obtained data [p. 51].

The data relevant to such a test, however, do not provide us with any esti-

mate of the congruity model's accuracy other than to tell us that the obtained values have moved some amount greater than zero in the direction of the predicted values—a modest result which could just as easily have been predicted by a much simpler “regression-toward-the-mean” hypothesis.

2. Osgood, Suci, and Tannenbaum (1957) state,

Another estimate of the accuracy of prediction is the average magnitude of deviation . . . between predicted and obtained scores for word combinations [p. 280].

They explicitly indicate, after looking at their data, that the congruity formula does not reliably predict the results obtained on the *evaluation* factor.¹⁰ There is therefore no difference between their combination study and our's insofar as the evaluative factor is concerned; neither set of data—their's or our's—is, considered as a whole, in accord with the congruity model. In this connection we calculated for our evaluation data the correlation between the size of the congruity principle's prediction errors and the magnitude of the disparity between the components and we obtained a rho of +.80. This suggests that the greater the disparity of the two interacting components the less apparent the operation of the congruity principle.

3. Osgood, Suci, and Tannenbaum (1957) write concerning the Osgood and Ferguson word combination study,

Still another, and perhaps the best, estimate of prediction accuracy is the correlation between predicted and obtained mean factor scores. . . . For the evaluative factor, $r = .86$; for the potency factor, $r = .86$;

¹⁰ But, they add, the congruity formula does reliably predict the results obtained on the potency and activity factors. Regarding these data it is highly likely that the disparity of component scores was highly restricted in their study (as it was in the present study) thus making it difficult, if not impossible, for the obtained values to deviate significantly from the predicted values.

and for the activity factor, $r = .90$ —all highly significant [p. 280].

In their attitude change study, Osgood and Tannenbaum (1955) report a correlation of .91 between predicted and obtained changes on the evaluation factor.

There are two reasons why we are reluctant to accept these findings as evidence of prediction accuracy. First, it is difficult to reconcile the admittedly negative findings with respect to the evaluation scores discussed in Item 2 above with the high correlations between obtained and predicted scores on the evaluative factor. Second, and far more important, while the high correlations seem at first glance to be impressive evidence for the congruity model, further reflection suggests that the correlations may be spurious. If the upper and lower values of the two components vary from one combination to the next, and if the obtained and predicted values of the combinations fall anywhere between the upper and lower values, a positive correlation will necessarily result even if the two sets of values (predicted and obtained) deviate significantly from one another. For example, the upper and lower limits on Assertion 3 (Table 2) is 6.52 for GOD and 5.38 for NEGRO. Table 1 shows that the predicted and obtained values for the configuration NEGRO *who believes in* GOD fall in between these two values (6.32 and 6.42, respectively). Similarly, the upper and lower limits on Assertion 8 are 5.84 for ATHLETE and 1.53 for DISHONEST, and the predicted and obtained values for DISHONEST ATHLETE again fall in between these values (3.36 and 1.80, respectively). When the two coordinates of Assertions 3 and 8 are plotted on a scatter diagram along with the co-ordinates for the remaining 10 assertions a high correlation is found ($r = .956$) even though the obtained values differ significantly from the con-

gruity model's predicted values in 8 of the 12 assertions. In other words, we too get high correlations between obtained and predicted values even though our results for the most part deviate markedly and significantly from those predicted by the congruity principle. It is thus seen that the correlation between the obtained and the predicted values cannot provide a test of the accuracy of prediction since the correlation is spuriously influenced by the particular upper and lower values of the two components entering into the combination, which inevitably delimit the range within which the predicted and obtained values of the combination will fall.

CONCLUDING REMARKS

Of all the major balance theories currently extant, the congruity principle is virtually the only one which, until now, has attempted to make specific quantitative predictions regarding the outcome of cognitive interaction. It thus put itself out on a limb where it could more easily be proven inadequate or insufficient. It is to the credit of the authors of the congruity principle that they themselves have drawn attention to the lack of precision in the predictions generated by the congruity model. It is also to their credit that they have tried to improve the predictive efficiency of the congruity model by positing a number of variables which limit its operation. These include a quantitative correction for credulity, an assertion constant and qualitative corrections for relevance-nonrelevance, derogation-nonderogation, and adjective-noun.

In our opinion, all these attempts at improving predictive efficiency stem from the fact that the basic theory on which it is based is psychologically an untenable one: congruity theory attempts to predict cognitive interaction solely from a knowledge of measurable

properties of the components judged in isolation. It is this deficiency which the present formulations are designed to overcome. We think it fair to say, from all the preceding, that despite the various attempts by Osgood and his co-workers to improve the congruity model's predictive efficiency, that even their own data do not provide support for the congruity principle, especially with regard to the evaluative factor, the most important factor in semantic differential studies and the one which is of major interest to attitude theory and research. Consequently, the problem of reconciling our present findings and interpretations, involving the principle of belief congruence, with those of Osgood's would seem to be largely resolved. The data presented here, designed to test the relative predictive efficiency of the congruity principle and the principle of belief congruence, would seem to be clearly more consistent with the latter principle. It remains to be seen whether this principle, as formulated earlier (Rokeach, 1960) and further elaborated here, will also prove helpful in studying other phenomena involving cognitive interaction, including attitude change.¹¹

¹¹ The problem of attitude change concerns the effects of cognitive interaction on changing the subsequent evaluative meaning of the components, or, those beliefs-disbeliefs activated by the components. While we have presented no data here on this issue it may be assumed, following Osgood and his co-workers, that the principle of belief congruence would predict effects on components, as a function of learning, similar to those predicted for the meaning of word combinations.

REFERENCES

- ASCH, S. E. Forming impressions of personality. *Journal of Abnormal and Social Psychology*, 1946, 41, 258-290.
- ASCH, S. E. *Social psychology*. New York: Prentice-Hall, 1952.
- BROWN, R. Models of attitude change. In, *New directions in psychology*. New York: Holt, 1962. Pp. 1-85.
- BYRNE, D., & WONG, T. J. Racial prejudice, interpersonal attraction, and assumed dissimilarity of attitudes. *Journal of Abnormal and Social Psychology*, 1962, 65, 246-253.
- EDWARDS, A. L. *Experimental design in psychological research*. New York: Holt, 1960.
- OSGOOD, C. E. Cognitive dynamics in human affairs. *Public Opinion Quarterly*, 1960, 24, 341-365.
- OSGOOD, C. E., SUCI, G. J., & TANNENBAUM, P. H. *The measurement of meaning*. Urbana: Univer. Illinois Press, 1957.
- OSGOOD, C. E., & TANNENBAUM, P. H. The principle of congruity in the prediction of attitude change. *Psychological Review*, 1955, 62, 42-55.
- ROKEACH, M. *The open and closed mind: Investigations into the nature of belief systems and personality systems*. New York: Basic Books, 1960.
- ROKEACH, M. *The three Christs of Ypsilanti: A psychological study*. New York: Knopf, 1964.
- ROKEACH, M. The nature of attitudes. In, *International encyclopedia of social sciences*. New York: Macmillan, 1965, in press.
- ROKEACH, M., SMITH, PATRICIA W., & EVANS, R. I. Two kinds of prejudice or one? In M. Rokeach, *The open and closed mind*. New York: Basic Books, 1960. Pp. 132-168.
- STEIN, D. D., HARDYCK, JANE A., & SMITH, M. B. Race and belief: An open and shut case. *Journal of Personality and Social Psychology*, 1965, 1, 281-289.

(Received February 18, 1964)

PERSON AND POPULATION AS PSYCHOMETRIC CONCEPTS¹

JANE LOEVINGER

Social Science Institute, Washington University

Person is the central primitive notion of psychology. Persons constitute populations, hence (or, that is to say), can be randomly sampled. There are no populations of tests, items, or testing conditions, hence, no means for random sampling. Sampling in those realms is almost invariably expert selection. Psychometric methods whose derivations assume random sampling of tests, items, or testing conditions include the Q technique, classical reliability theory and its recent liberalizations, and others. These techniques seem inadequately anchored in relation either to the science of psychology or to the practice of testing. Representativeness of experimental conditions as the foundation for generalization is reaffirmed, however. Expert attention to representativeness is required precisely when randomness is unattainable.

The two problems that are basic to psychometrics and differential psychology are the logic of measurement and the factorial problem. Under what conditions can one be said to measure a psychological trait in some more or less rigorous meaning of measurement? And, what are the basic dimensions of individual differences among people? Thus, how can we measure? And, what is worth measuring? These two questions are the framework of this essay.

My thesis is that person and test are radically different kinds of concepts and that confusion or exchange of them leads at best to ad hoc pseudomeasurements and at worst to error. Person is a primitive, undefinable term (Strawson, 1958); surely it is the central primitive notion of psychology, though rarely acknowledged as such. Population is a collective form, and so is

sample, though less obviously so etymologically. Test, on the other hand, refers to an instrument or artifact of our science and contains an implication of measurement or indexing. Our samples comprise persons; tests are analyzed statistically, e.g., they are correlated and factored. Statistical or psychometric methods that purport to sample tests or items seem to violate these most basic notions of psychological statistics. But such a position stands in opposition to parts of the work of such leading psychometricians as Burt (1937); Cattell (1952); Stephenson (1936, 1952); Guttman (1944, 1950); Coombs (1952); Block (1961); Lord (1955); Tryon (1957); and Cronbach, Rajaratnam, and Gleser (1963).

Often tied to sampling of items is correlation of persons, i.e., correlation of a set of scores generated by or pertaining to one person with other corresponding sets of scores. Correlation of persons also may arise in other contexts. No attempt will be made here to list or to evaluate all such possibilities.

¹Preparation of this paper was supported by Grant MH-05115 from the National Institute of Mental Health, United States Public Health Service. It is a revision of the Presidential address to Division 5 of the American Psychological Association, Philadelphia, September 2, 1963.

Some authors, such as Coombs, have stressed an abstract equivalence, hence interchangeability, of person and test. Others, such as Guttman, have used an imprecise concept of a "universe of attributes," from which the content of particular tests has been derived by a sampling process, also somewhat imprecise, though explicitly not a random process (Guttman, 1950b, p. 54). The imprecision has been criticized (Campbell & Kerckhoff, 1957; Loevinger, 1955) in a manner anticipating the present article. Recently a number of psychometricians, particularly Lord (1955 and elsewhere) and Cronbach, Rajaratnam, and Gleser (1963 and elsewhere), have rendered precise the notion of sampling items and tests and capitalized on it to derive psychometric formulas from statistical theory. In doing so, they have also opened the way for clearer and more precise evaluation of the notion of random sampling of items and tests.

TEST VERSUS PERSON

Every schoolboy knows what a *test* is. He might be surprised to know that the first, oldest meaning given by the dictionary is

a cupel or cupeling hearth for refining precious metals,

and the next,

examination or trial by the cupel; hence, any critical examination or decisive trial.

The educational use corresponds to that of the schoolboy:

Any series of questions or exercises or other means of measuring the skill, knowledge, intelligence, capacities, or aptitudes of an individual or group.

The meaning that seems closest to the general meaning of the term in psychology is the following:

Means of trial; specif., subjection to conditions that show the real character of person or thing in a certain particular; as, the

tuberculin test for tuberculosis [*Webster's New Collegiate Dictionary*, 1961, p. 878].

The schoolboy meaning, in which test stands for a set of items, differs from that of psychometricians, who use the term rather for the items, administration, scoring key, and norms. Plagiarism of each other's items tends to be condoned by psychometricians because of the importance of norms, which can hardly be plagiarized. These connotations of the word test are not arbitrary, but reflect the realities of the situation; the score of a person by no means depends on the items alone.

In regard to the concept of person, on the other hand, every schoolboy does know what the word means, in the same sense that psychometricians use it. No one could be enrolled in kindergarten who did not have the concept of person as a deeply ingrained part of his mentality, who did not understand and recognize the identity of a person from day to day despite superficial changes in appearance, and who did not know the difference between encountering one person and encountering two persons. A person retains his identity through the extreme changes from infancy to old age. In contrast with a person, a test does not have an identity through time despite superficial changes in administration or norms, nor is it always clear whether one is confronted with two copies of a single test or with two different tests.

Seemingly minor differences in test administration may greatly affect performance on many tests; hence, the use of norms set for one administrative circumstance is inappropriate if the administration changes, say, if time limits are altered. Since it requires different norms, it is not the same test. A test may change even though there is no change in items, administration, or norms. This fact is recognized and

dealt with by periodic revisions, in some cases by yearly revisions.

It has been argued (Loevinger, 1957) that test and retest are best considered as two different tests except where they are shown to be closely similar with respect to mean, variance, and correlation with outside criteria. In the usual case, where secular trends exist, i.e., where test and retest have an appreciable mean difference, different norms are mandatory; hence, by definition, test and retest are different tests. The argument here is slightly different: it is ambiguous whether test and retest are the same or two different tests. The ambiguity over identity differentiates the concept of test from that of person.

There are, thus, drastic differences between person as a construct and test as a construct. A person retains recognizable identity through slight superficial changes, while a test may or may not. We are always clear whether we are confronted by one person or by two people. People do not, in front of one, shade off into one another imperceptibly. Tests, on the other hand, may differ in minor ways whose significance is doubtful. Are two forms of the same test a single test or two different ones? Considering the limiting form of a test, a single item, how great does the change in wording have to be before it becomes a new item? Some changes in wording may be ignored by most people (e.g., a small misspelling or a grammatical error), hence be negligible changes. Others may greatly affect the difficulty or popularity of the item. In general, one cannot predict in advance, but must try out both forms.

Some psychologists who have practiced or advocated correlation of persons or sampling of tests or testing conditions have acknowledged arguments like those presented here; others

have ignored the issues. Burt (1937, and in earlier writings) probably was the first to exploit systematically the correlations between persons, though he credits W. Stern with the original proposal. Thinking primarily of persons serving as rating instruments, Burt makes the point that what is in one context a person is in another context a test. Hammond (1955) has observed that this situation is typical of much clinical research. Burt (1937) goes on:

We can safely correlate *persons* only when the traits or tests fulfill much the same conditions that we should require of the persons when correlating *tests*: that is to say, as a general rule: (i) the tests must be sufficiently numerous to keep the probable errors low; (ii) they must form a representative or random sample of the total population of tests; (iii) the marks or measurements obtained from different tests must all be reducible to terms of the same unit; (iv) the distribution of the measurements should be, as nearly as possible, normal; and (v) the correlation should be linear. Otherwise a correlation between persons is likely to be meaningless [p. 68].

To the caveat of Burt must be added the following, from Cattell (1952):

The test population, except for the personality sphere concept, does not have the qualities of a biological species population. The experimenter, indeed, usually does not even make any attempt to sample the test universe [p. 510].

Cattell's case for the personality sphere having the character of a biological species population rests on the claim that personality traits can be counted on to be represented in language, hence a list of trait names covers the personality domain (Cattell, 1946).

To Burt's caution and Cattell's stricture may be appended some remarks of Spearman (1927), made in slightly different context. He argued against the definition of intelligence as either an average or a sample of functions on the grounds that to compute an

average or to draw a sample certain basic postulates must be met that are not met in intelligence tests: unequivocal domain, comparable cases, no repetition, and no omission. The limits of the domain over which sampling or averaging takes place must be defined unequivocally, a single case or instance must be in some respect equivalent to every other case or instance, and the domain must be defined so as neither to omit nor to repeat instances.

The worst about this theory of sampling, however, has yet to be mentioned. It is that, in actual practice, the procuring of a genuine sample has not really even been attempted. One of the many indications of this is the prevalent procedure, in the construction of a series of tests, of trying out a large number and then selecting those which exhibit the highest correlations with all the rest. Such a procedure seems to have been more or less influential, directly or indirectly, in the framing of all generally accredited series at the present day. . . . In order to obtain a genuine sample, one carefully retains, and even adds to, the sorts which were at first little represented, and which therefore tend to be *least* correlated with, the remainder [p. 70].

As we shall see, Spearman was describing a kind of representative, not random, sampling. Returning to Cattell's assertion that the personality sphere constitutes a population, one sees that his argument rests entirely on fulfilling the postulate of no omission. He makes no claim of unequivocal domain, of comparable cases, or of no repetition.

POPULATION AND SAMPLE

In examining the basic constructs of psychometrics more deeply, let us begin where statistics began, with the concept of *population*. What is usually considered the first statistical tract was "Natural and Political Observations Made upon the Bills of Mortality" by Graunt, published in 1662 (Graunt, 1956). Graunt noted that parish

clerks of London published weekly an account of deaths, listed by causes, and also of christenings. He collated these for the years 1604-61 and showed how inferences could be drawn from the data. Ages of the casualties were not given in the reports, but Graunt made some estimates of age from the causes of death. The number dying of acute diseases he took as a measure of the tendency of the climate and air to sudden unfavorable changes; the number dying of chronic diseases was a measure of the fitness of a country for longevity, etc.

The first life insurance tables were published in 1693 by Sir Edmund Halley. He used tables of births and funerals in the city of Breslaw as a source for mortality tables from which he calculated the worth of annuities taken out at various ages. For this purpose, he noted, the bills of mortality of London and similar data from Dublin were not adequate:

First, In that the *Number* of the People was wanting. Secondly, That the *Ages* of the People dying was not to be had. And Lastly, That both *London* and *Dublin* by reason of the great and casual Accession of *Strangers* who die therein, (as appeared in both, by the great Excess of the *Funerals* above the *Births*) rendered them incapable of being Standards for this purpose; which requires, if it were possible, that the People we treat of should not at all be changed, but die where they were born, without any Adventitious Increase from Abroad, or Decay by Migration elsewhere [Halley, 1956, p. 1437].

Here is the statistical notion of population *in statu nascendi*. Population meant people, and at that time only persons. Obviously a somewhat broader usage prevails today. Tippett (1943) states:

I shall refer only to populations consisting of recognizably discrete individuals, e.g., men or electric lamps [p. 79].

In his usage only denumerable classes of objects constitute populations.

In current statistical usage most of the populations referred to are hypothetical ones; usage has shifted to referring to the measurements as the population rather than the objects to which the measurements pertain, but the latter point is not germane to the present discussion. Wallis and Roberts (1962) state:

The two most fundamental concepts of statistics are those of a *sample* and a *population*. A *sample* is often referred to as "the data" or "the observations": numbers that have been observed. The *population*, on the other hand, is the totality of all possible observations of the same kind [pp. 124-125].

And again:

Population is an abstract concept fundamental to statistics. It refers to the totality of numbers that would result from indefinitely many repetitions of the same process of selecting objects, measuring or classifying them, and recording the results. A population is, thus, a fixed body of numbers, and it is this general body of numbers about which we would like to know [pp. 139-140].

In current statistical usage the notion of population has been greatly generalized from its popular and original meaning. In regard to the concept of *sample*, however, the statistical usage is derived from ordinary usage not by generalization but by a narrowing of meaning. Thus a dictionary definition of sample is:

A part of anything presented for inspection, or shown as evidence of the quality of the whole; a specimen [*Webster's New Collegiate Dictionary*, 1961, p. 748].

But statistical inference can give information only about random samples. A section of the Wallis and Roberts book (pp. 143-146) is devoted to "Randomness vs. Expert Selection." They do not disparage experts or expert judgment but emphasize that statistical methods are applicable only to random not to expert selection.

SAMPLING OF TESTS AND ITEMS

Have any psychological tests been marketed whose items are chosen by a random sampling method? Is it not rather the case that all serious test construction projects involve expert selection? Here is an enormous discrepancy. In one building are the test and subject matter experts doing the best they can to make the best possible tests, while in a building across the street the psychometric theoreticians construct test theories on the assumption that items are chosen by random sampling.

The objection to all psychometric developments that assume random sampling of items or tests is in the first instance that they grossly misrepresent the actual case, which is almost invariably expert selection rather than random sampling. But there is also implied in the argument a subtler and deeper point. The term population implies that in principle one can catalog, or display, or index all possible members, even though the population is infinite and the catalogue cannot be completed. Statistical sampling must be tied to such a display and indexing system, else it cannot be random. Tests and items are not that sort of thing. There is no meaning to talking about populations of tests or items. No system is conceivable by which an index of all possible tests could be drawn up; there is no generating principle. Any definition of the terms population and sample, however abstract, that retains those elements of meaning that authorize the use of statistical theory will exclude tests and test items.

Q-SORT TECHNIQUE

Another group of psychometric usages that violate fundamental meanings includes *Q* technique and *Q* sort. Take as a single and particularly meritorious

example Block's (1961) CQ set. This is a set of 100 descriptive phrases by means of which professional persons make a standardized description of a person, patient, class of persons, etc. The sorter arranges the statements in a given quasi-normal distribution. Ordinal numbers are assigned to the successive, arbitrary intervals. Each person rated has 100 scores, ranging in value from 1 to 9, one for each statement sorted. The 100 items are eclectic, in some cases semitechnical, but theoretically noncommittal. As a more or less standardized means for rendering descriptions of patients and other persons mutually comparable, the CQ set has obvious merit.

The two fundamental problems of psychometrics and differential psychology are evaded, however. The number assigned to each item for each person rated comes from an arbitrary distribution. The question of how nonarbitrary numbers, i.e., rigorous measures, can be constructed is not only not solved, it is foreclosed from view. The method may work in a rough and ready way, but it cannot contribute to understanding the fundamental nature of psychological measurement.

The other basic problem, the factorial problem, is also in part evaded. The items are not advanced on the basis of, nor do they contribute to, any fundamental insights into the structure of human personality. How then can the method work even as an ad hoc descriptive instrument? Adequate coverage of the domain of individual differences was obtained not by sampling a population of items, for no such population exists, but by an unacknowledged and informal sampling of persons. Block reports that his set of items was modified more than once after use and criticism by a number of professional colleagues of varying interests. There are two instances of

sampling here. First there is a sample of colleagues; then there is the second order sample, consisting of patients, friends, and relations of each colleague. Each professional person presumably asked himself, in effect, can I describe the distinctive features of the patients and other people I know in terms of these items? If not, what further items are necessary to make the list complete? Thus, the merits of the CQ set as a sample of traits derive from an informal sampling of people. In this case the nonrandomness of the sampling is of little consequence since what is at stake is something more like a range than a central tendency.

RELIABILITY AND ITS REVISIONS

Cronbach, Rajaratnam, and Gleser (1963) have reworked test theory in terms of generalizability. In administering a test one is not interested in the performance on that test per se, but in generalizing to something else. All kinds of reliability and validity can be construed as species of generalization, and they have done a service in urging care in the design and interpretation of reliability studies to support the appropriate generalizations. But they are led to base their formulations not only on random sampling of persons but also on random sampling of items and of testing conditions. The same considerations hold with regard to test conditions as with regard to items. Everyone who has a serious investment in applied testing seeks optimal and not random testing conditions. There are unknown and uncontrollable aspects of the test situation, but one works always towards optimization not randomization.

The question of how to generalize to performances other than those sampled and to circumstances other than those of the test situation is not trivial.

What is meant by the best items and the optimal testing conditions is, in part at least, those that best mediate the most predictions.

Cronbach, Rajaratnam, and Gleser present random sampling of items as a less restrictive alternative to the assumption of the equivalence of parallel forms in classical reliability theory. But the assumption of parallel tests must be based either on random sampling of a hypothetical population of parallel tests or else on random sampling within strata of items; in some formulations each item constitutes a stratum. Thus, if random sampling of tests and items is an untenable assumption, classical theory of reliability as well as the revisions of Lord (1955), Tryon (1957), and Cronbach, Rajaratnam, and Gleser (1963) must be basically modified.

Lord (1955 and elsewhere) uses the assumption of random sampling of items without examination of its implications as an assumption. Tryon (1957) presents an interesting historical review of the assumptions used in classical derivations of reliability formulas. He shows that two assumptions have been used, one based on division of scores into hypothetical true scores and error factors, which he calls the Spearman-Yule assumption, and the other that all items have equal standard deviations and equal intercorrelations, which he calls the Brown-Kelley assumption. Most authors have used one or the other, a few have used both. Neither, however, is necessary. He defines the reliability coefficient as the correlation between a test and a hypothetical comparable test composed of the same number of items, having the same average item variability and the same interitem correlation. Clearly implied is that the items are from the same behavior domain, since the other conditions could be satisfied by totally

uncorrelated tests. In the simplest case, where the items of a test form an unstratified composite, the items "are as if drawn at random from a large pool" of items (Tryon, 1957, p. 233). From this it follows that the mean covariance of items in the two forms will equal the mean covariance of items within each form, which is crucial to Tryon's derivations. The assumption of parallel tests, which Cronbach, Rajaratnam, and Gleser are concerned to replace, is a direct consequence of the Brown-Kelley assumption, or, if one begins with the Spearman-Yule assumption, is needed to bring the estimates of true score and error factor within reach.

In the present view, then, the attempts by Tryon and by Cronbach, Rajaratnam, and Gleser to derive reliability formulas on less restrictive assumptions than have been used in the past have not been fundamental enough. The modified assumption still involves sampling of items. A previous criticism of classical reliability, that reliability cannot be defined in a noncircular manner (Loevinger, 1947, 1957), is another, and perhaps more searching, way of looking at the same basic problem. To show how the present considerations can be related to the critique of reliability in terms of circularity, consider Tryon's system. Tryon assumes that tests comprise items randomly drawn from a single "behavior domain." How then does one establish, as the test constructor must, that two items are drawn from the same behavior domain? There can hardly be any answer other than the one usually given in textbooks: the items within a single domain are intercorrelated to the extent of their reliabilities. Here reliability is used to define the behavior domain, just as the behavior domain construct is used to define reliability.

BRUNSWIK'S PRINCIPLE OF REPRESENTATIVE DESIGN

Brunswik (1947) advocated substituting representative for systematic design of psychological experiments. He emphasized that before making psychological generalizations one must consider how far the conditions of the supportive experiments represent the organism's natural ecology.

Representative sampling is extended from the subjects to the objects, from the individuals to the stimulus situations and tests [Brunswik, 1947 p. 56].

At first his position seems to be directly contrary to that of the present paper. However, he did not consider precisely the present point, i.e., the difference between expert selection of a representative sample and truly random sampling. Statistical sampling theory applies only to random sampling, but a judgment of representativeness underlies all probabilistic generalization. The question of whether something like random sampling can be assumed to have occurred must be judged separately in every different context.

In commenting on studies of social perception where people serve both as subjects (in effect, as tests) and as objects of study, Brunswik noted that the subjects are usually numerous, and a correlative attention is paid to the representativeness of the sampling. Concerning the objects of social perception, he showed that in many studies they were too few, and that little attention had been paid to their representativeness of any population. In clinical research the situation noted by Brunswik is most often reversed: the persons studied may be numerous and chosen with some care, whereas the clinicians serving as measuring instruments are usually few and unrepresentative of any definable class of clinicians or people. Brunswik's plea, that research

workers in social perception attend to the representativeness of persons serving as objects, needs to be turned around in clinical research. To protect the generality of the findings more attention should be paid to the representativeness of people serving as subjects.

The more one pursues this line of thought, the more arbitrary it seems what one calls object and what one calls subject; and the more one comes to Burt's and, later, Hammond's point, that what is from one view a person is from another view a test. Apparently, whatever the role of persons in the experimental design, whether as subject or object, notions of random sampling are appropriate in that dimension, and probably a serious attempt at something like random sampling should occur far more frequently than it now does.

Several authors have raised the problem of generalizing over testing conditions, and this again is reminiscent of Brunswik's discussion of the narrowness of the environmental sample that enters most psychological studies. It is difficult to think of instances where the principle of ecological representativeness would be of value in relation to the conditions of testing; one essentially always seeks optimal testing conditions.

In relation to the content of psychological tests, greater representation of the variety of situations of everyday life in place of the often artificial restriction of subject matter seems much needed. Rarely, however, will there be possible anything like a random sampling of life situations. The importance of expert attention to the representativeness of the content is emphasized rather than lessened by abandoning the random sampling model.

Goodenough (1949) pointed out that tests may be construed as signs or as

samples of behavior; Loevinger (1957) later expanded the point, showing that tests are always and necessarily samples of behavior as well as signs of non-test behavior. But to say that they are samples is not to say that they are random samples; moreover, they are always evaluated in terms of how well they serve as signs of nontest behavior. Brunswik's term, representativeness, is ambiguous here; it could be taken to mean randomness, or, perhaps more logically, sign value. Any vagueness or ambiguity obscures the point of the present essay.

Randomness is a property not of an individual sample but of the process of sampling. . . . A sample of size n is said to be a *random sample* if it was obtained by a process which gave each possible combination of n items in the population the same chance of being the sample actually drawn. . . . Nonstatisticians usually assume that the importance of randomness arises from the "fairness" and lack of bias with which such samples represent the population. This is important, of course, but of more importance is the fact that *the pattern of sampling variability for any population is known if, but only if, the sampling is random* [Wallis & Roberts, 1962, pp. 141-142].

Brunswik's point, the necessity of representative design as a foundation for generalization, a point elaborated in relation to a number of significant psychometric problems by Cronbach, Rajaratnam, and Gleser, neither requires nor justifies assumption of random sampling.

RASCH'S PSYCHOMETRICS

Rasch (1960) has devised a truly new approach to psychometric problems, although one of somewhat limited applicability. He makes use of none of classical psychometrics, but rather applies algebra anew to a probabilistic model. The probability that a person will answer an item correctly is as-

sumed to be the product of an ability parameter pertaining only to the person and a difficulty parameter pertaining only to the item. Beyond specifying one person as the standard of ability and one item as the standard of difficulty, the ability assigned to an individual is independent of that of other members of the group and of the particular items with which he is tested; similarly for the item difficulty. The model of course is not universally applicable. Indeed, these two properties were once suggested as criteria for absolute scaling (Loevinger, 1947); at that time proposed schemes for absolute scaling had not been shown to satisfy the criteria, nor does Guttman scaling do so. Thus, Rasch must be credited with an outstanding contribution to one of the two central psychometric problems, the achievement of nonarbitrary measures.

Note that Rasch is concerned with a different and more rigorous kind of generalization than Cronbach, Rajaratnam, and Gleser. When his model fits, the results are independent of the sample of persons and of the particular items within some broad limits. Within these limits, generality is, one might say, complete. Of course, one only studies particular items and particular people, and the problem of generalizing comes up again in setting the limits within which the model holds. The people can be considered a random sample from some population, but the items cannot. Rasch suggests that the model be turned around and be used to select items that fit it. If this is done, the goodness of fit must be evaluated with a new sample of people. The question of what traits the given items yield information about is part of the factorial problem; in relation to this problem Rasch's current methods have not enlarged our understanding.

PERSON AS A PRIMITIVE NOTION

Strawson (1958, 1959) has examined philosophically how man constructs and apprehends his universe. Our frame of reference is a spatio-temporal one; within this framework material bodies are the basic particulars, i.e., the particulars that can be identified, both distinguished and re-identified, without reference to any other kind of particulars.

Private experiences have often been the most favoured candidates for the status of "basic" particulars; on the present criteria, they are the most obviously inadmissible [Strawson, 1959, p. 41].

The category of persons is basic in a different but related way.

The admission of this category as primitive and underived [is] a necessary condition of our membership of a non-solipsistic world [Strawson, 1959, p. 246].

In acquiring the idea of a person, we learn to assign certain predicates to people, to ourselves at the same time as to others. In particular, the meaning of the predicate *to be depressed* is the conjunction of the behaviors by which we assign it to others and the feelings by which we assign it to ourselves.

The dictionaries do not give two sets of meanings for every expression which describes a state of consciousness: a first-person meaning and a second-and-third-person meaning [Strawson, 1959, p. 99].

On purely logical grounds Strawson has thus arrived at a position similar to that which Hebb (1960) supports on psychological grounds, i.e., that we learn about ourselves to a considerable extent the same way and at the same time that we learn about others. The position is consonant with the developmental findings of Piaget.

Let us take off from Strawson's position as a starting point. Tests, since they are processes rather than

material bodies, are certainly not basic particulars, as persons are. Tests are rather a means of assigning predicates to persons; they are "conditions that show the real character of a person . . . in a certain particular." Tests are instruments of psychologists and as such are arbitrary. Presumably, however, there are nonarbitrary long-lasting differences among persons to which the relations among tests provide clues. That is precisely the factorial problem. The factorial problem has been a focal one for many of our greatest psychometricians—Spearman, T. L. Kelley, and Thurstone.

The factorial problem in its classic form springs directly from the concepts of person and test in their received meanings. Proposals to sample items or tests and some of the schemes to subject persons to statistical analysis are made in defiance of accepted usage. How can so many able psychometricians be so wrong? Or to state the matter noncommittally, though I am not uncommitted, how can trained people differ by so much as I differ from what seems to be a majority of my colleagues? An answer comes from Polanyi, who, prior to becoming a social scientist, was a physician and later one of the leading physical chemists of the world; he thus brings to social science an intimate knowledge of the very arts and sciences that the scientifically rigorous among us most desire to emulate. Polanyi (1958) decries the aspiration of many psychologists and other social scientists to construct a completely objective discipline, to renounce any commitment as an unscientific embarrassment, and to depersonalize their science. The differences among psychometricians can be described in his terms.

I am committed to the meaning of person as it occurs in psychology and in everyday life; to the meaning of test

as it is used, e.g., by the American Psychological Association Committee on Tests; and to the meanings of sample and population as they are used in statistics. Those who use methods exchanging persons and tests seem to me to be saying: "It doesn't matter what we call a person and what we call a test. The formula simply calls for a set of commensurate scores. If we are completely objective, what is the difference whether it is the score of several people on one test or of one person on several tests? How do we know whether we are correlating persons or tests? How do we even know whether the person has failed the test or the test has failed the person? If we are objective, how do we know whether we are sampling persons or sampling items?"

For myself, I know. My position is both more objective and less objective than theirs: more objective in the sense that I believe that there are objective realities independent of my mind that exist out there and that I as a scientist essay to comprehend; less objective in the sense that I acknowledge a personal commitment to what is meaningful and important. I am not indifferent to whether statistics are applied in a meaningful or meaningless way. On the basis of my commitments I know that it is meaningful to sample persons and not to sample tests, that persons pass or fail tests, but tests do not pass or fail persons; i.e., persons, not tests, make up populations, and that persons are variables only insofar as they serve as tests.

Polanyi argues that a scientist who acknowledges his personal commitment has taken a self-consistent position, while to deny one's commitment is to take an intrinsically contradictory position. In what follows his arguments will be applied to psychometrics, which he did not consider. In saying it does

not matter whether we enter tests where the problem calls for persons and persons for tests, the psychometrician is saying, in effect, "I am completely objective, hence uncommitted to the received meaning of the words person and test." But having taken this stand on the strength of noncommitment, he is then committed to those usages (for me, distorted ones) of statistics that result from exchange of persons and tests. For ultimately no scientist can work and remain uncommitted. The moment he enters his laboratory or sits down to his typewriter he is making a commitment. Each of us faces a short tenure and a vast universe. What we choose to work on we are declaring thereby meaningful and significant, in fact the most significant thing that we can do.

Polanyi (1958) summarizes his "fiduciary programme" in the statement,

I believe that in spite of the hazards involved, I am called upon to search for the truth and state my findings [p. 299].

This is not a subjective or solipstic position. It is an affirmation of an independently existing and meaningful reality.

To be sure, science has often advanced by breaking out of old meanings and opening new possibilities, as Einstein did when he redefined simultaneity. Obviously, population has also been progressively redefined. At first it meant a class of people; then a class of denumerable objects. Perhaps what it now means is a class of objects or events, usually hypothetical, that can be randomly sampled ("the totality of numbers that would result from indefinitely many repetitions of the same process of selecting objects"). Thus the burden of the definition is shifted to the term *random*.

PATTERN RECOGNITION

Randomness is, like person, almost or quite undefinable, but recognizable by what Polanyi calls connoisseurship. In any essay "The Logic of Biology" Grene (1961), following Polanyi, shows that recognition of pattern is the basis of all scientific activity, especially scientific theories. She goes on, in biology there is a refinement of the ability to detect patterns that forms the basis for recognition of species. No entirely objective criterion has been evolved to supplant the naturalist's ability to recognize species, usually instantaneously. Similarly, in everyday life we instantly recognize individual persons whom we know. Some larger animals are also recognized as individuals. The identification of persons as a form of pattern recognition and the identification of randomness as absence of pattern are thus instances of an indispensable but often unacknowledged ability of a scientist. A similar argument is made by Lorenz (1959).

The favorite program of the ultra-objective psychometrician, the computer search for patterns, is a caricature of the pattern-search of a great scientist. No computer could have found the pattern expressed in Newton's theory of universal gravitation, in the theory of relativity, in thermodynamics, or in quantum theory.

Does random sampling of tests, items, and testing conditions represent an illuminating new insight, or is it a meaningless gimmick, a quantitative trick without profound meaning for psychology and psychometrics? When such a question falls within his domain, every scientist has the responsibility, whether he acknowledges it or not, to apply his own pattern-finding ability to it. Crucial points are the difference between random and expert selection and the problem of continuing identity

in sampled objects. There is little hope of touching those who are deeply committed to the idea that to be scientific means being completely objective, uncommitted, detached, and approaching all issues in a depersonalized manner. Others who have used the methods here criticized will be reluctant to embrace the implications of these methods.

REFERENCES

- BLOCK, J. *The Q-sort method in personality assessment and psychiatric research*. Springfield, Ill.: Charles C Thomas, 1961.
- BRUNSWIK, E. *Systematic and representative design of psychological experiments*. Berkeley: Univer. California Press, 1947.
- BURT, C. Correlations between persons. *British Journal of Psychology*, 1937, 28, 59-96.
- CAMPBELL, E. Q., & KERCKHOFF, A. C. A critique of the concept "universe of attributes." *Public Opinion Quarterly*, 1957, 21, 295-303.
- CATTELL, R. B. *Description and measurement of personality*. Yonkers-on-Hudson: World Book, 1946.
- CATTELL, R. B. The three basic factor-analytic research designs: Their interrelations and derivatives. *Psychological Bulletin*, 1952, 49, 499-520.
- COOMBS, C. H. *A theory of psychological scaling*. Ann Arbor: University of Michigan, Engineering Research Institute, 1952.
- CRONBACH, L. J., RAJARATNAM, NAGESWARI, & GLESER, GOLDINE C. Theory of generalizability: A liberalization of reliability theory. *British Journal of Statistical Psychology*, 1963, 16, 137-163.
- GOODENOUGH, FLORENCE L. *Mental testing*. New York: Rinehart, 1949.
- GRAUNT, J. Natural and political observations made upon the bills of mortality. (Orig. publ. 1662) In J. R. Newman (Ed.), *The world of mathematics*. New York: Simon & Schuster, 1956. Pp. 1421-1435.
- GRENE, MARJORIE. The logic of biology. In, *The logic of personal knowledge*. Glencoe, Ill.: Free Press, 1961. Pp. 191-205.
- GUTTMAN, L. A basis for scaling qualitative data. *American Sociological Review*, 1944, 9, 139-150.
- GUTTMAN, L. The basis for scalogram analysis. In S. A. Stouffer (Ed.), *Meas-*

- urement and prediction. Princeton: Princeton Univer. Press, 1950. Ch. 3. (a)
- GUTTMAN, L. The problem of attitude and opinion measurement. In S. A. Stouffer (Ed.), *Measurement and prediction*. Princeton: Princeton Univer. Press, 1950. Ch. 2. (b)
- HALLEY, E. An estimate of the degrees of the morality of mankind, drawn from curious tables of the births and funerals at the city of Breslaw; with an attempt to ascertain the price of annuities upon lives. (Orig. publ. 1693) In J. R. Newman (Ed.), *The world of mathematics*. New York: Simon & Schuster, 1956. Pp. 1437-1447.
- HAMMOND, K. R. Probabilistic functioning and the clinical method. *Psychological Review*, 1955, 62, 255-262.
- HEBB, D. O. The American revolution. *American Psychologist*, 1960, 15, 735-745.
- LOEVINGER, JANE. A systematic approach to the construction and evaluation of tests of ability. *Psychological Monographs*, 1947, 61 (4, Whole No. 285).
- LOEVINGER, JANE. The universe. *American Psychologist*, 1955, 10, 399. (Abstract)
- LOEVINGER, JANE. Objective tests as instruments of psychological theory. *Psychological Reports*, 1957, 3, 635-694.
- LORD, F. M. Sampling fluctuations resulting from the sampling of test items. *Psychometrika*, 1955, 20, 1-22.
- LORENZ, K. Gestaltwahrnehmung als Quelle wissenschaftlicher Erkenntnis. [Gestalt perception as fundamental to scientific knowledge.] *Zeitschrift für experimentelle und angewandte Psychologie*, 1959, 6, 118-165. (Reprinted: In L. von Bertalanffy & A. Rapoport (Eds.), *General systems: Yearbook of the Society for General Systems Research*. Vol. 7. New York: Society for General Systems Research, 1962. Pp. 37-56.)
- POLANYI, M. *Personal knowledge*. Chicago: Univer. Chicago Press, 1958.
- RASCH, G. *Probabilistic models for some intelligence and attainment tests*. *Studies in Mathematical Psychology I*. Copenhagen: Danmarks Paedagogiske Institut, 1960.
- SPEARMAN, C. *The abilities of man*. New York: Macmillan, 1927.
- STEPHENSON, W. The foundations of psychometry; Four factor systems. *Psychometrika*, 1936, 1, 195-209.
- STEPHENSON, W. Some observations on *Q* technique. *Psychological Bulletin*, 1952, 49, 483-498.
- STRAWSON, P. F. Persons. In H. Feigl, M. Scriven, & G. Maxwell (Eds.), *Minnesota studies in the philosophy of science*. Vol. 2. *Concepts, theories, and the mind-body problem*. Minneapolis: Univer. Minnesota Press, 1958. Pp. 330-353.
- STRAWSON, P. F. *Individuals: An essay in descriptive metaphysics*. London: Methuen, 1959.
- TIPPETT, L. H. C. *Statistics*. London: Oxford Univer. Press, 1943.
- TRYON, R. C. Reliability and behavior domain validity: Reformulation and historical critique. *Psychological Bulletin*, 1957, 54, 229-249.
- WALLIS, W. A., & ROBERTS, H. V. *The nature of statistics*. New York: Collier, 1962.
- Webster's New Collegiate Dictionary*. Springfield, Mass.: G. & C. Merriam, 1961.

(Received March 23, 1964)

THEORETICAL NOTES

SOME METHODOLOGICAL PROBLEMS IN CATTELL'S MULTIPLE ABSTRACT VARIANCE ANALYSIS¹

JOHN C. LOEHLIN

University of Nebraska

Certain methodological problems in Cattell's Multiple Abstract Variance Analysis (MAVA) technique for the analysis of trait variance into hereditary and environmental components are considered. These include inconsistencies between and within some of the equations in the model and an apparent anomaly in correcting for errors of measurement. A new approach to the derivation of MAVA equations is developed, and implications for previously published MAVA results briefly discussed.

Over a number of years, Cattell and his associates (Cattell, 1960, 1963; Cattell, Blewett, & Beloff, 1955; Cattell, Kristy, & Stice, 1957) have been developing a method for human population genetics known as Multiple Abstract Variance Analysis (MAVA). This method breaks down observed trait variance into various components representing the contributions of heredity and environment to variation in the trait.

Essentially, the MAVA procedure is to test various groups of subjects (twins, siblings, etc.) reared together or apart and to infer from observed trait variances in these groups certain theoretical hereditary and environmental variances and covariances. This is done by writing sets of equations, in each of which an observed variance is expressed as a combination of theoretical variances and covariances, and then solving the equations simultaneously to obtain the theoretical values. These can then be used to calculate various nature-nurture ratios, and the like.

In principle, the MAVA method is a powerful one and should be able to yield important information about genetic and

environmental influences on traits. It contains within it the more traditional twin and foster-sib methods and hence provides all the information they can, plus the additional information derivable from considering them jointly. However, the MAVA method in its current form appears to suffer from certain difficulties, and it is the purpose of the present paper to examine some of these briefly. Two major kinds of difficulties will be considered. The first involves internal inconsistencies in the MAVA equations as they are presented by Cattell in his most extensive theoretical treatment of the method (Cattell, 1960). The second is an apparent systematic error in previous applications of the equations to data by Cattell and his co-workers (Cattell et al., 1955; Cattell et al., 1957). Along with the examination of these difficulties, a new approach to deriving MAVA-type equations will be developed, and implications for previously published MAVA results considered.

INCONSISTENCIES IN THE MAVA EQUATIONS

Two types of inconsistencies within the system of MAVA equations will be discussed in this section. The first is inconsistency between equations dealing with within- and between-family variance; the second, inconsistency in the handling of selective placement effects in the equations involving adopted children.

¹ The author wishes to thank R. B. Cattell for pointing out several obscurities in an earlier version of this paper, and encouraging a more systematic critique of the Multiple Abstract Variance Analysis model. The author would also like to thank K. W. Schaie and F. J. Dudek for their comments on the present version.

Within- and Between-Family Variances

In order to examine particular discrepancies in the Cattell equations, it will be useful to develop a general form of the MAVA equation, from which the particular equations for twins, sibs, etc., may be derived. (For simplicity, the equation will be developed for the usual case involving pairs of individuals.)

Following the general rationale of the MAVA method, it is assumed that the deviations of the true scores of a pair of individuals from the population mean of some attribute may be broken into additive components, as follows:

$$\begin{aligned} x_1 &= x_{wh1} + x_{we1} + x_{bh1} + x_{be1} \\ x_2 &= x_{wh2} + x_{we2} + x_{bh2} + x_{be2} \end{aligned} \quad [1]$$

where the x 's represent deviation scores; wh , we , bh , and be refer, respectively, to within-family heredity, within-family environment, between-family heredity, and between-family environment; and 1 and 2 designate the two individuals in the pair.

Now twice² the total variance of the trait in the population may be divided into two additive portions, σ_w^2 and σ_b^2 , the former representing variation within pairs of individuals, and the latter representing variation between pairs. The former can be calculated as:

$$\sigma_w^2 = \frac{1}{2N} \sum_1^N (x_1 - x_2)^2, \quad [2]$$

and the latter as:

$$\sigma_b^2 = \frac{1}{2N} \sum_1^N (x_1 + x_2)^2; \quad [3]$$

in each case N represents the number of pairs. Substituting in Equation 2 the expressions for x_1 and x_2 from Equation 1, we obtain:

$$\begin{aligned} \sigma_w^2 &= \frac{1}{2N} \sum_1^N (x_{wh1} + x_{we1} + x_{bh1} + x_{be1} \\ &\quad - x_{wh2} - x_{we2} - x_{bh2} - x_{be2})^2; \end{aligned}$$

² This is to bring the present development into line with that of Cattell, who calculates σ_w^2 and σ_b^2 as separate estimates of the population variance, with $\sigma_w^2 + \sigma_b^2$ thus corresponding to twice σ_x^2 in the general population (Cattell et al., 1955, p. 126).

expanding, we have:

$$\begin{aligned} \sigma_w^2 &= \frac{1}{2N} \sum_1^N (x_{wh1}^2 + x_{we1}^2 + \cdots + x_{be2}^2 \\ &\quad + 2x_{wh1}x_{we1} + 2x_{wh1}x_{bh1} + \cdots \\ &\quad + 2x_{bh2}x_{be2}); \end{aligned}$$

expressing as variances and covariances:

$$\begin{aligned} \sigma_w^2 &= \frac{1}{2}\sigma_{wh1}^2 + \frac{1}{2}\sigma_{we1}^2 + \cdots + \frac{1}{2}\sigma_{be2}^2 \\ &\quad + r_{wh1we1}\sigma_{wh1}\sigma_{we1} + r_{wh1bh1}\sigma_{wh1}\sigma_{bh1} + \cdots \\ &\quad + r_{bh2be2}\sigma_{bh2}\sigma_{be2}; \end{aligned}$$

assuming that corresponding 1s and 2s are equal in variance and covariance (since assignment as 1 or 2 is purely arbitrary), and collecting terms:

$$\begin{aligned} \sigma_w^2 &= \sigma_{wh}^2(1 - r_{wh1wh2}) + \sigma_{we}^2(1 - r_{we1we2}) \\ &\quad + \sigma_{bh}^2(1 - r_{bh1bh2}) + \sigma_{be}^2(1 - r_{be1be2}) \\ &\quad + 2\sigma_{wh}\sigma_{we}(r_{wh1we1} - r_{wh1we2}) \\ &\quad + 2\sigma_{wh}\sigma_{bh}(r_{wh1bh1} - r_{wh1bh2}) \\ &\quad + 2\sigma_{wh}\sigma_{be}(r_{wh1be1} - r_{wh1be2}) \\ &\quad + 2\sigma_{we}\sigma_{bh}(r_{we1bh1} - r_{we1bh2}) \\ &\quad + 2\sigma_{we}\sigma_{be}(r_{we1be1} - r_{we1be2}) \\ &\quad + 2\sigma_{bh}\sigma_{be}(r_{bh1be1} - r_{bh1be2}). \end{aligned} \quad [4]$$

Now by assuming particular values of the correlations in parentheses, this general equation may be reduced to any one of Cattell's within-family MAVA equations. For example, consider sibs reared together in the same family (see Equation 4 in Table 1, Cattell, 1960):

$$\sigma_{ST}^2 = \sigma_{wh}^2 + \sigma_{we}^2 + 2r_{whwe}\sigma_{wh}\sigma_{we}.$$

This equation can be derived from Equation 4 above by assuming that: (a) the correlations r_{wh1wh2} and r_{we1we2} are zero—each individual receives an independent assortment of the parental genes and family influences; (b) the correlations r_{bh1bh2} and r_{be1be2} are unity—the two members come from the same family; (c) r_{wh1we1} does not, in general, equal r_{wh1we2} (see comment below); and (d) in all other cases, the two correlations within the parentheses are equal. Actually this last follows automatically from Assumption 2—the two r 's must be equal in each case because one variable is shared and the other correlated 1.00. Assumption 3 brings out a point not obvious in Cattell's version of the equation, that his r_{whwe}

represents a *difference* between two correlations: the tendency for a given hereditary characteristic to select or attract certain environmental influences ($r_{wh_1we_1}$), and the fact that the genetically influenced traits of one pair member serve as part of the environment to which the other is subjected ($r_{wh_1we_2}$). Thus cases could arise where this term is negligible even though both correlations are appreciable.

A general equation for between-family variance can be derived in the same manner as the above, starting from Equation 3. This works out to be:

$$\begin{aligned}\sigma_b^2 = & \sigma_{wh}^2(1 + r_{wh_1wh_2}) + \sigma_{we}^2(1 + r_{we_1we_2}) \\ & + \sigma_{bh}^2(1 + r_{bh_1bh_2}) + \sigma_{be}^2(1 + r_{be_1be_2}) \\ & + 2\sigma_{wh}\sigma_{we}(r_{wh_1we_1} + r_{wh_1we_2}) \\ & + 2\sigma_{wh}\sigma_{bh}(r_{wh_1bh_1} + r_{wh_1bh_2}) \\ & + 2\sigma_{wh}\sigma_{be}(r_{wh_1be_1} + r_{wh_1be_2}) \\ & + 2\sigma_{we}\sigma_{bh}(r_{we_1bh_1} + r_{we_1bh_2}) \\ & + 2\sigma_{we}\sigma_{be}(r_{we_1be_1} + r_{we_1be_2}) \\ & + 2\sigma_{bh}\sigma_{be}(r_{bh_1be_1} + r_{bh_1be_2}). \quad [5]\end{aligned}$$

It will be noted that this is identical to Equation 4 for σ_w^2 , except that the signs within parentheses are all positive rather than negative. It will also be seen that

$$\begin{aligned}\sigma_w^2 + \sigma_b^2 = & 2\sigma_{wh}^2 + 2\sigma_{we}^2 + 2\sigma_{bh}^2 + 2\sigma_{be}^2 \\ & + 4r_{wh_1we_1}\sigma_{wh}\sigma_{we} + 4r_{wh_1bh_1}\sigma_{wh}\sigma_{bh} \\ & + 4r_{wh_1be_1}\sigma_{wh}\sigma_{be} + 4r_{we_1bh_1}\sigma_{we}\sigma_{bh} \\ & + 4r_{we_1be_1}\sigma_{we}\sigma_{be} + 4r_{bh_1be_1}\sigma_{bh}\sigma_{be},\end{aligned}$$

and that this is equal to $2\sigma_z^2$.

As in the within-family case, we can by suitable assumptions about the r 's derive Cattell's various equations. For example, the between-family equation for ordinary sibs reared together (see Equation 10, Table 1, Cattell, 1960) is:

$$\begin{aligned}\sigma_{BNF}^2 = & \sigma_{wh}^2 + \sigma_{we}^2 + 2\sigma_{bh}^2 + 2\sigma_{be}^2 \\ & + 2r_{whwe}\sigma_{wh}\sigma_{we} + 4r_{bhbe}\sigma_{bh}\sigma_{be}.\end{aligned}$$

To derive this from Equation 5 requires the same Assumptions 1, 2, and 3, as for the within-family case, plus the stronger assumption that all the other r 's are zero except $r_{bh_1be_1}$ and $r_{bh_1be_2}$, which are equal, since $r_{be_1be_2} = 1.00$. (Note that if the r_{whwe} here and in Cattell's within-family equation are the same, $r_{wh_1we_2}$ must also be zero.)

Proceeding in this fashion to examine the rest of Cattell's equations, it soon becomes evident that there are quite a few instances in which the assumptions required to derive the within-family equation are different from, and sometimes in contradiction with, those required for the between-family equation.³ For example, Cattell's equations for identical twins reared together (Equations 1 and 8, Table 1, Cattell, 1960) are:

$$\begin{aligned}\sigma_{ITT}^2 = & \sigma_{we'}^2 \\ \sigma_{BITTF}^2 = & 2\sigma_{wh}^2 + \sigma_{we}^2 + 2\sigma_{bh}^2 + 2\sigma_{be}^2 \\ & + 2\sqrt{2}r_{whwe}\sigma_{wh}\sigma_{we} + 4r_{bhbe}\sigma_{bh}\sigma_{be}.\end{aligned}$$

The first inconsistency involves the σ_{we}^2 terms, where $r_{we_1we_2}$ must be assumed to be zero to obtain the between-family equation from Equation 5, but where the special $\sigma_{we'}^2$ term in the within-family equation allows the twins' environments to be positively correlated—i.e., $\sigma_{we'}^2 = \sigma_{we}^2(1 - r_{we_1we_2})$. The second inconsistency involves the within-family heredity-environment correlation: to derive the within-family equation, one must assume that $r_{wh_1we_1}$ and $r_{wh_1we_2}$ are equal; to derive the between-family equation, one must assume that they are not. (The former assumption is presumably the correct one, following as it does from $r_{wh_1wh_2} = 1.00$ and $r_{wh_1we_1} = r_{wh_2we_2}$, and would lead to the fifth term in the equation for σ_{BITTF}^2 becoming $4r_{whwe}\sigma_{wh}\sigma_{we}$.)

This last discrepancy in particular has major consequences for the solutions presented in Cattell's (1960) paper, as it bears directly on the possibility of separating hereditary and correlated environmental variance.

In his initial presentation of the MAVA method, Cattell did not solve separately and directly for the effects of heredity within families (the variance σ_{wh}^2) and correlated environmental influences (the covariance $r_{whwe}\sigma_{wh}\sigma_{we}$). He used instead the indirect approach of setting r_{whwe} to various values and choosing the value that gave the most satis-

³ Such contradictions can be found in the equations for identical twins reared together, sibs reared apart, half-sibs reared together, half-sibs reared apart, and unrelated children reared together.

factory results on other grounds—mainly in bringing σ_{wh}^2 close to σ_{bh}^2 . The reason for the difficulty in the separate solution for σ_{wh}^2 and r_{whwe} is clear when one considers how r_{whwe} arises. Heredity and environment may become correlated within families if infants differing in genes receive systematically different environmental inputs—whether this is due to differing reactions of others to genetically based differences in physique or temperament, or due to the infant's own selection of his environment (cf. Cattell, 1963). How does one separate such covariation from more direct effects of the genes on the trait in question? Presumably, by finding groups of subjects of similar genetic diversity which differ systematically in the environmental pressures brought to bear on the traits, or in opportunities for selection of favorable or unfavorable trait environments. While such groups might be found, it does not appear that the groups the present MAVA method uses have been selected to vary systematically along such dimensions. The difficulty in separating σ_{wh}^2 and r_{whwe} in the present MAVA design is therefore understandable. In his 1960 paper, however, Cattell presents direct solutions for σ_{wh}^2 and r_{whwe} —which is puzzling, since the data used are still the same in relevant respects. An examination of the sets of equations suggests, however, that the solution is an artifact resulting from the discrepancy noted above. This may be most easily seen in Cattell's Table 2 (Cattell, 1960) presenting the 10 equations of his "Limited Resources Design." Inspection of this table reveals that Equation 1 contains neither σ_{wh}^2 nor $r_{whwe}\sigma_{wh}\sigma_{we}$, and that in all but one of the other equations this variance and covariance stand in a constant ratio to each other: a solution is thus algebraically possible only because of the divergent equation, Number 7—and this is the between-family equation for identical twins discussed above. If the indicated correction of the covariance term to $4r_{whwe}\sigma_{wh}\sigma_{we}$ is made, this equation falls into line with the rest, and separate solution for σ_{wh}^2 and r_{whwe} becomes algebraically impossible. Hence it appears that if these variables are to be separated,

it must be by incorporating appropriate additional groups in the design.

Alternatively, it might be proposed that one estimate the within-family values from those between families; but this raises the question whether analogous problems attend the separation of heredity and correlated variance in the between-family case. The covariance in this instance involves two components: in part, covariation may result from influences existing prior to the birth of the subjects in question—systematic relationships may exist between parental genes and family environmental conditions; and in part covariation may arise, as in the within-family case, from environmental response to the subject or selection of the environment by the subject. By the same arguments as in the within-family case, it would appear that the MAVA design in its present form will be unable to separate the second kind of heredity-environment covariance from direct hereditary effects; the design should, however, be capable of estimating correlations due to effects of the first kind.

A final comment may be appropriate concerning the relative advantages of using within- and between-family variance. Ordinarily, if there is a difference, the within-family equation will be simpler than the corresponding between-family equation for a given set of assumptions about r 's. While Equations 4 and 5 involve exactly the same variables, under typical sets of simplifying assumptions all the terms that vanish from Equation 5 also vanish from Equation 4, and frequently more, tending to make it easier to work with the within-family equations. An additional argument for using the within-family variances is that they are probably better behaved in the face of sampling inadequacies: if foster families, say, come from a restricted sector of the population, this will certainly make the observed between-family variance suspect, but the within-family data may still be usable.

Effects of Selective Placement

Besides the inconsistencies noted above between different equations, there are

some cases in the current MAVA model where inconsistent assumptions are made within equations. An instance is the treatment of selective placement effects.

One of the refinements Cattell introduced into his 1960 version of MAVA was a separate $r_{bhb'e}$ to allow for the possibility of heredity and environment becoming correlated by selective placement of infants by adoption agencies. However, the effect of such placement on the variances is not taken into account. Consider the case of sibs reared apart in separate families. A placement-agency $r_{bhb'e}$ is introduced into the between-family equation (see Equation 11, Table 1, Cattell, 1960) to allow for the possibility that the placement agency may be influenced in its placements by knowledge concerning the child's biological family. Suppose, for example, that the placement agency attempts to place each sib in a family like the one from which he came. To the extent that the agency succeeds, this will affect $r_{bhb'e}$, as indicated. But it will also affect the environmental variance term: the sib pairs will wind up in less different families than if they had been randomly assigned, and σ_{be}^2 should be decreased by an appropriate factor within pairs and increased between pairs. Similar considerations apply in the equations for identical twins reared apart and unrelated children reared together. While in all these cases one may argue whether a placement correlation is needed, clearly if it is to be included placement effects must be allowed for in the variances as well.

CORRECTION FOR ERRORS OF MEASUREMENT

In his discussions of the MAVA model, Cattell has argued persuasively for the appropriateness of removing error of measurement variance from the observed variances prior to entering the theoretical equations. In the actual application of the MAVA procedure to empirical data, however, there appears to have been a systematic error. Consider the basic measurement operation, measuring a trait in an individual. The measurement error involved should be roughly the

same whether measuring the trait in an identical twin or in a member of an unrelated pair. And indeed this appears to have been the case in the empirical data: the reliabilities quoted in Cattell, Blewett, and Beloff (1955) and in Cattell, Kristy, and Stice (1957) fluctuate from group to group, as would be expected from the small sample sizes, but show no systematic tendency to decrease across groups. And yet Cattell's corrections do: the error variance removed is systematically less when going from unrelated individuals to sibs to identical twins. This appears to have resulted from the use of inappropriate reliabilities in formulas of the type $(1 - r_{IT})\sigma_{IT}^2$ in calculating the error variances. By this, the error variance for a group (in this case identical twins) is obtained by multiplying the within-pair variance by the complement of the reliability; the reliabilities used, however, were not those of the differences from which the within-pair variances were calculated, but those of the individual test scores. Now the reliabilities of differences will not in general be the same as those of the individual scores, being less as the correlation between scores increases (Mosier, 1951, p. 777); hence, use of the original score reliabilities in formulas like the above will tend to systematically underestimate the error variance for all but the general population pairs. This underestimation will be progressively greater as the pair correlation increases, being greatest for the identical twins. One can perhaps see this most clearly by considering a hypothetical case where the "pair" consists of two tests of the same individual. Here the "within-pair" variance should obviously all be error variance—but this result will not follow if the test reliability is substituted in the expression above. However, the reliability of a difference score is zero in this case and thus yields a correct result.⁴

⁴ One need not, in practice, go to all this trouble—it can be shown that the use of Mosier's difference-score reliability formula in Cattell's expression is equivalent to estimating the error variance directly from $(1 - r_u)$, where r_u is the test reliability. The test reliability may be obtained either for the

The effect of the undercorrection of the correlated groups will depend on the particular way the observed variances enter into the MAVA equations. In general, the variance within pairs of identical twins will be relatively too large (tending to inflate estimates of within-family environmental variance), and differences between the variances of different groups will be too small. Thus in the 1955 paper the hereditary variance between families (σ_{bh}^2), which was obtained by the subtraction $\sigma_{UT}^2 - \sigma_{ST}^2$, was consistently underestimated, leading to puzzling difficulties in its relation to σ_{wh}^2 : according to the expectations of the authors based on genetic theory (Cattell et al., 1955, p. 137), σ_{bh}^2 should have been equal to or somewhat larger than σ_{wh}^2 , but it persisted in being markedly smaller. The proper error correction tends to make for a larger σ_{bh}^2 and a more satisfactory relationship between the two.

An attempt was made to systematically rework Cattell's published data using appropriate error corrections, but unfortunately the small N s and low test reliabilities resulted in fairly erratic data—a number of the corrected variances

total population or the specific subgroup in question, depending on one's theoretical preference (see Cattell et al., 1955, for a discussion).

were negative—and the effort was abandoned. It is the writer's impression, however, that on the whole the effect of the corrections was to increase estimates of between-family variance relative to within-family variance, and of heredity relative to environment, compared to Cattell's published estimates.

REFERENCES

- CATTELL, R. B. The multiple abstract variance analysis equations and solutions. *Psychological Review*, 1960, **67**, 353-372.
- CATTELL, R. B. The interaction of hereditary and environmental influences. *British Journal of Statistical Psychology*, 1963, **16**(Part 2), 191-210.
- CATTELL, R. B., BLEWETT, D. B., & BELOFF, J. R. The inheritance of personality: A multiple variance analysis of approximate nature-nurture ratios for primary personality factors in Q-data. *American Journal of Human Genetics*, 1955, **7**, 122-146.
- CATTELL, R. B., KRISTY, N. F., & STICE, G. F. A first approximation to nature-nurture ratios for eleven primary personality factors in objective tests. *Journal of Abnormal and Social Psychology*, 1957, **54**, 143-159.
- MOSIER, C. I. Batteries and profiles. In E. F. Lindquist (Ed.), *Educational measurement*. Washington, D. C.: American Council on Education, 1951. Pp. 764-808.

(Received March 17, 1964)

ON THE ASSUMPTION OF A STEEPER AVOIDANCE GRADIENT IN MILLER'S CONFLICT THEORY

RICHARD S. BOGARTZ

Institute of Child Behavior and Development, University of Iowa

Miller (1959) derives the prediction that in an approach-avoidance situation the animal sometimes will approach the goal, but stop short of it. He assumes 2 gradients which cross, and he represents these gradients by straight lines. He then asserts that the deductions could be made on the basis of any curves that have a continuous negative slope that is steeper for avoidance than approach at each point above the abscissa. This assertion is correct; however, it is shown here that the assumption of an everywhere steeper avoidance gradient rules out exponential decay curves as gradients and, in fact, rules out all monotonically decreasing, differentiable functions which have a limiting value of 0 as distance from the goal increases. Requiring that the gradients cross appears sufficient.

Miller (1959) has presented a theoretical treatment of approach-avoidance conflict in which he derives the prediction that under certain conditions the animal will tend to approach the goal when started far from it, but will stop approaching before the goal is reached. He assumes the existence of an approach gradient and an avoidance gradient which cross, such that near the goal the avoidance gradient is higher than the approach gradient, but far from the goal this relation is reversed. It is assumed that the animal approaches the goal if the approach gradient is higher than the avoidance gradient and avoids it if the avoidance gradient is higher than the approach gradient. It follows, then, that the animal tends "to remain in the region where the two gradients intersect [p. 206]."

In the graphic representation of this derivation (see his Figure 1), Miller uses straight-line gradients, but he remarks,

It is only for the sake of simplicity that the gradients are represented by straight lines in these diagrams. Similar deductions could be made on the basis of any curves that have a continuous negative slope that is steeper for avoidance than for approach at each point above the abscissa.

This statement is correct, however certain functions, e.g., those appearing in Figures 10 and 11 of Miller's (1948) paper

on displacement, do not satisfy these requirements.

It will be shown here that in the situation where the gradients are assumed to cross, the requirement that the avoidance gradient have a steeper slope than the approach gradient at each point above the abscissa is incompatible with a large class of functions. In particular, if it is assumed (a) that each gradient is a continuous, monotonically decreasing function of distance from the goal, which is differentiable at least once at every distance greater than zero, and (b) that the limiting value of each function is zero, then it cannot be assumed that the avoidance gradient is everywhere steeper than the approach gradient.

Let x be the distance from the goal; let $F(x)$ and $G(x)$ be the avoidance and approach gradients, respectively. Assume the following.

Assumption 1 (monotonic decreasing gradients): if $x_2 > x_1$, then $F(x_2) < F(x_1)$ and $G(x_2) < G(x_1)$.

Assumption 2 (limiting value of each gradient is zero):

$$\lim_{x \rightarrow \infty} F(x) = \lim_{x \rightarrow \infty} G(x) = 0.$$

Assumption 3 (steeper avoidance gradient at all points): the first derivatives of $F(x)$ and $G(x)$, $F'(x)$ and $G'(x)$, exist for all values of $x > 0$; and, for any x , $F'(x) < G'(x)$.

Since the proof is only for the situation in which the gradients cross, the features of this situation must also be specified.

Crossing of gradients condition: there exists a point x_0 such that (a) $F(x_0) = G(x_0)$; (b) if $x < x_0$, $F(x) > G(x)$; and (c) if $x > x_0$, $F(x) < G(x)$.

Consider two points x_1 and x_2 such that $x_0 < x_1 < x_2$. From elementary calculus it is known that

$$F(x_2) - F(x_1) = \int_{x_1}^{x_2} F'(x) dx$$

and

$$G(x_2) - G(x_1) = \int_{x_1}^{x_2} G'(x) dx$$

Therefore, by Assumption 3,

$$\begin{aligned} G(x_2) - G(x_1) &= \int_{x_1}^{x_2} G'(x) dx \\ &> \int_{x_1}^{x_2} F'(x) dx = F(x_2) - F(x_1). \end{aligned}$$

Hence, by addition and subtraction,

$$F(x_1) - G(x_1) > F(x_2) - G(x_2)$$

By Part c of the crossing-of-gradients condition, the value on the left-hand side of this inequality is a negative number. However, by Assumption 2, the value on the right-hand side (also negative) may be brought as close to zero as desired by selecting x_2 sufficiently large. In particular, it can be brought closer to zero than is $F(x_1) - G(x_1)$. But this contradicts the inequality and completes the proof.

Exponential decay curves exemplify the problem. Let $F(x) = ae^{-ix}$ and $G(x) = be^{-jx}$, where $a > b$ and $i > j$.

The functions satisfy Assumptions 1 and 2 and the crossing-of-gradients condition. Assumption 3 requires that $-iae^{-ix} = F'(x) < G'(x) = -jbe^{-jx}$, or, rearranging terms, that $ia/jb > e^{(i-j)x}$. Obviously a sufficiently large value of x can be chosen such that the inequality is false.

No special brief is held here for either the monotonicity assumption or the zero-limit assumption. The former seems psychologically reasonable and consistent with Miller's (1959) first two postulates (p. 205). The latter also seems reasonable in that one might expect that both the approach and the avoidance gradients could be made as close to zero as desired by placing the animal sufficiently far from the goal; however, it is not clear to what extent this assumption is included in Miller's thinking. The point to be made here is that the requirement of an everywhere steeper avoidance gradient appears to be unnecessarily severe. The crossing-of-gradients condition seems to suffice for the predictions with which Miller is concerned.

REFERENCES

- MILLER, N. E. Theory and experiment relating psychoanalytic displacement to stimulus-response generalization. *Journal of Abnormal and Social Psychology*, 1948, 43, 155-178.
- MILLER, N. E. Liberalization of basic S-R concepts: Extensions to conflict behavior, motivation, and social learning. In S. Koch (Ed.), *Psychology: A study of a science*. Vol. 2. *General systematic formulations, learning, and special processes*. New York: McGraw-Hill, 1959. Pp. 196-292.

(Received August 7, 1964)

IN DEFENSE OF REMOTE ASSOCIATIONS

KENT M. DALLETT

University of California, Los Angeles

Slamecka (1964) has argued that the evidence for remote associations resulting from the use of the association method is of dubious validity, owing to a response-bias artifact. Slamecka's argument fails to be convincing: on both logical and experimental grounds it appears that the response-bias artifact does not invalidate the results of the association method.

In order to support his contention that the doctrine of remote associations (RAs) is of "doubtful validity" Slamecka (1964) has tried to show that each of the methods commonly used to demonstrate RAs does so by means of an artifact. The purpose of this note is to point out that Slamecka's attack on the association method does not, as he supposes, cast doubt upon the existence of RAs as demonstrated by that method.

In the association method the subjects learn a serial list and are then shown items from the list, one by one in scrambled order, with instructions to give the first other item from the list that comes to mind. If the subject gives the item which, in learning, followed the item presented, his response is classified as an adjacent forward association. If he gives any other association, his response is classed as remote and assigned a degree of remoteness depending upon its linear distance from the stimulus syllable in the original list order. Since the presence of responses other than the originally "correct" one could be accounted for on several grounds, users of the association method typically grant no importance to the simple fact that such RAs occur and focus instead upon the fact that the frequency of RAs decreases as their degree of remoteness increases. This will be referred to as the RA gradient, and the existence of such a gradient will be taken as evidence that the RA doctrine is still, in some sense, a useful one. However, the raw RA gradient which results from simple tabulation of RAs is unacceptable evidence, because there are more opportunities for RAs of low degrees of re-

moteness to occur than there are opportunities for RAs of high degrees of remoteness to occur. A "correction for opportunity" has been traditional, in which the frequencies of RAs with few opportunities to occur are multiplied by an appropriate factor: the best-known example of such a correction is that given by Raskin and Cook (1937). Slamecka (1964) claimed that an additional artifact must be taken into account. According to Slamecka, subjects may be biased to give items from near the ends of the list as responses, as a result of the relatively rapid learning of end items usually described as the serial position effect. Given the conditions of the association method, and a bias to give end items as responses Slamecka (1964) states,

Such a state of affairs should logically, in and of itself, produce the pattern of results commonly reported [p. 69].

However, Slamecka does not explain the logical necessities involved, and in fact it seems more likely that Slamecka's bias should *not* "logically, in and of itself" produce the RA gradient. The reason for this is that end items, but not middle items, can enter into RAs of high degrees of remoteness—hence a bias in favor of end items should flatten or invert the usual RA gradient, while a bias favoring middle items should enhance the gradient. Slamecka tried to counterfeit an RA gradient by giving the subjects the appropriate response bias in the absence of serial learning and supported the argument given here: after correction for opportunity he found virtually no forward gradient, and an *inverted* gradient of

TABLE 1
EXAMPLE OF RA DATA, TAKEN FROM SLAMECKA'S (1964) TABLE 2,
AFTER REMOVAL OF ADJACENT FORWARD ASSOCIATIONS

Stimulus items	Response items						Total
	1	2	3	4	5	6	
(N)	—	11	7	4	8	9	39
1	—	—	8	2	9	14	33
2	24	—	—	4	8	11	47
3	14	9	—	—	14	11	48
4	13	14	10	—	—	8	45
5	26	8	6	5	—	—	45
6	17	13	12	6	6	—	54
Total	94	55	43	21	45	53	311

backward RAs. Hence we may flatly reject Slamecka's (1964) conclusion that

the overall results of the association method can be brought about without any previous serial learning [p. 71].

In deciding that his artifactual gradient successfully approximated an RA gradient, Slamecka appears to have depended upon its similarity in shape to the U-shaped gradients reported by Raskin and Cook (1937). A closer examination of Slamecka's data, however, suggests that the similarity is not great enough to be convincing, and further suggests that the troublesome U shape in the Raskin and Cook data has been somewhat over-emphasized by the correction for opportunity usually used in such studies.

The major drawback of the conventional correction for opportunity is that it requires multiplying small (and possibly unreliable) frequencies of infrequent RAs by large correction factors. The result is that minor differences in RA frequencies are likely to become unduly magnified. An excellent example appears in the data of Raskin and Cook (1937), in which RA frequencies of 12 and 11 become 42 and 77, respectively, after correction. An unwary reader is likely to be taken in by (a) the large difference between the corrected values, or (b) the large values themselves; although in actual fact the data reveal a difference of one item between frequencies which can-

not possibly exceed their expected values by more than 11 or 12—probably less, depending on the nature of the expected values. One can design a better correction, involving a comparison of RA gradients obtained in an experiment, with the gradients to be expected if differential opportunity alone were operating. In designing such a correction method, there are two further sources of bias which must be considered. First, the limited opportunities for RAs of high degrees of remoteness are further restricted by the fact that a subject may give more non-remote or "correct" responses to the easier end items (which participate in RAs of high degrees of remoteness). Such a bias is not taken into account by the usual correction for opportunity, which assumes that every item serves equally often as the stimulus for an RA. The second bias is along the lines indicated by Slamecka: if the subjects use responses from the end of the list more often than they use responses from the center of the list, they are more likely to give RAs of high degrees of remoteness and vice versa.

For convenience in illustrating the proposed correction method, the data of Slamecka's (1964) Experiment IV will be used. In Table 1 a matrix of the obtained RA frequencies appears. The rows represent stimulus items, and the columns represent the same items used as responses. The cell entries are the

frequencies with which a given response occurred to a given stimulus in the association test. Slamecka placed a separate stimulus item in initial position (X); in most RA data this does not appear. Prior to analysis, it will be noted that nonremote (adjacent forward) associations have been removed. In some analyses (e.g., Dallett, 1959), an association in which the first item in the list is given in response to the last item is considered an adjacent forward association and is also removed, leaving the lower left-hand cell vacant.

Using the appropriate marginal frequencies, an expected matrix is constructed, incorporating the observed stimulus bias, response bias, or both. For example, the matrix of expected RAs, taking only stimulus bias into account, would have entries of 7.80 in each of the nonvacant cells of the first row, under the assumption that of the 39 RAs elicited by Stimulus (X), Responses 2, 3, 4, 5, and 6 are equally likely, and therefore each cell entry is $39 \div 5$. The entries in each nonvacant cell of the second row would be $33 \div 4$ or 8.25. Referring to Table 1, it will be seen that the *obtained* frequency for forward RAs of degree 1 is $11 + 8 + 4 + 14 + 8 = 45$. The expected frequency, on the basis of stimulus

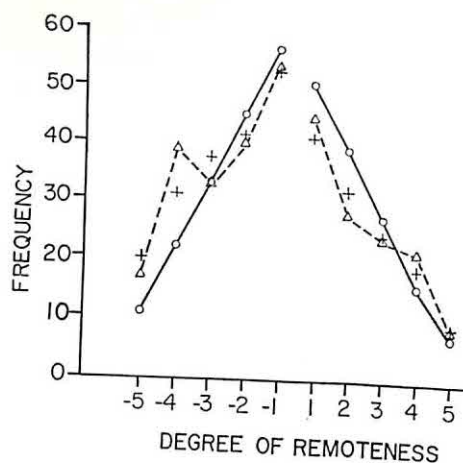


FIG. 1. Expected (circles) and obtained (triangles) RA gradients from Slamecka's Experiment IV. (The crosses indicate the effect of an additional correction for response bias applied to the expected values.)

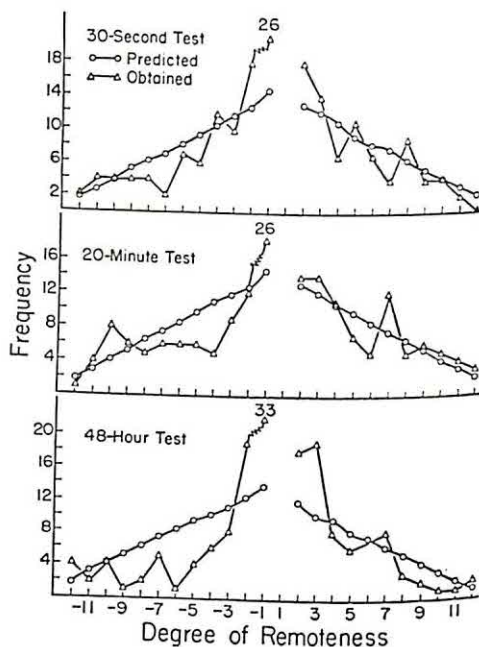


FIG. 2. A comparison of expected ("predicted") and obtained RA gradients for the subjects tested at different intervals after learning (Dallett, 1959).

bias and opportunity alone, is $7.80 + 8.25 + 11.75 + 12.00 + 11.25 = 51.05$. It will be noted that the correction for opportunity arises from the fact that both expected and obtained matrices contain the same disproportionate opportunities for RAs of low degrees of remoteness to occur. For forward RAs of Degree 5, there is only one opportunity, with an obtained frequency of 9 and a predicted frequency of 7.80.

Figure 1 presents the obtained and expected RA gradients for Slamecka's data.¹ The expected gradients are represented in

¹ The conventions for numbering degrees of remoteness need clarification. In the discussion of Slamecka's data, adjacent forward associations are labeled "degree 0" in agreement with Slamecka's usage. In Figure 1, however, adjacent *backward* associations appear as "degree -1"; Slamecka calls these "degree 0" also. In Figure 2, the convention adopted was to call the *stimulus* "degree 0"; hence, the first backward degree of remoteness is "degree -1," and adjacent forward associations (not plotted) are of "degree +1."

circles with solid lines, while the obtained gradients appear as triangles and dashed lines. As previously indicated, Slamecka's obtained RA gradient appears to be *flatter* than one would expect on the basis of opportunity and stimulus bias alone. It should be noted that the mean difference between expected and obtained gradients must be zero, so that valid interpretations are limited to the different shapes of the gradients, and different meanings cannot be given to deviations below and deviations above the expected gradient.

An analysis incorporating these stimulus-bias and opportunity corrections was carried out in an earlier study (Dallett, 1959), using actual data obtained by the association method. Expected (circles) and obtained (triangles) RA data from that study appear in Figure 2. The association test was given immediately, 20 minutes, or 48 hours, after the subjects learned a list of 14 CVCs to a 79% criterion of mastery. Inspection of Figure 2 reveals that the obtained RA gradient is in general steeper than the expected one. A correction for response bias was not attempted, but would probably have very little effect, the frequencies of usage of Responses 1-14 being, respectively, 40, 35, 39, 49, 31, 25, 30, 45, 34, 37, 31, 50, 49, 44. A serial position response bias is not apparent in these data.

It is possible to incorporate an approxi-

mate correction for response bias as well. The expected matrix for Slamecka's data, with both stimulus and response bias incorporated, appears in Table 2. To illustrate the computation of cell entries, the first entry in the first row is $(55 \times 39) \div 217$; in other words, the row total is distributed among cells in such a way that cell entries are proportional to the relevant column totals. The reason for this being an approximate correction is that each row total is associated with *different* columns, and it will be noted that the column totals are not recovered exactly, as they would be, say, in a contingency table for which complete independence of rows and columns was a tenable hypothesis. Despite the approximate nature of the correction, it will be noted from the crosses in Figure 1, that the expected RA gradient, taking response bias into account, is an improved predictor for 9/10 data points, as would be expected if we attribute the flatness of Slamecka's RA gradient to response bias effects.

Slamecka's data and the data of Dallett (1959), when considered together, indicate that the gradient of RAs which the association method reveals is not entirely artifactual, although the methods of analysis usually used have ignored some artifacts or possible artifacts associated with serial position effects. A subject does not respond randomly when en-

TABLE 2

MATRIX OF EXPECTED RA FREQUENCIES, BASED UPON THE MARGINAL TOTALS OF TABLE 1: BOTH STIMULUS BIAS AND RESPONSE BIAS ARE TAKEN INTO ACCOUNT

Stimulus items	Response items						Total
	1	2	3	4	5	6	
(X)	—	9.88	7.73	3.77	8.09	9.52	38.99
1	—	—	8.75	4.28	9.17	10.80	33.00
2	20.74	—	—	4.63	9.93	11.69	46.99
3	18.26	10.68	—	—	8.74	10.30	47.98
4	17.26	10.10	7.89	—	—	9.73	44.98
5	19.86	11.61	9.08	4.44	—	—	44.99
6	19.67	11.51	9.00	4.39	9.42	—	53.99
Total	95.79	53.78	42.45	21.51	45.35	52.04	310.92

couraged to respond freely. Whether this is a basis or byproduct of serial learning and the serial position effect is still an open question.

REFERENCES

DALLETT, K. M. Retention of remote associations. *Journal of Experimental Psychology*, 1959, 58, 252-255.

RASKIN, E., & COOK, S. W. The strength and direction of associations formed in the learning of nonsense syllables. *Journal of Experimental Psychology*, 1937, 20, 381-395.

SLAMECKA, N. J. An inquiry into the doctrine of remote associations. *Psychological Review*, 1964, 71, 61-76.

(Received April 13, 1964)

IN DEFENSE OF REMOTE ASSOCIATIONS

B. R. BUGELSKI

State University of New York at Buffalo

Recently Slamecka criticized the concept of remote associations and attacked 3 kinds of evidence for the concept. His attacks are analyzed and questioned as to relevance to the issue. A 1-trial learning experiment is reported wherein familiar items are shown to follow the traditional pattern of distribution of remote associations. Experimental results refute deductions from Slamecka's arguments. It is concluded that not only are direct verbal associations formed in serial learning (contrary to Slamecka's claim) but that remote associations are obviously formed in early stages of learning.

The logical and experimental attack by Slamecka (1964) on the concept of remote associations (RAs) is a remarkable tour de force. His analysis of three lines of evidence for such associations led him to the conclusion that not only are RAs nonexistent, but that in the learning of serial lists even direct associations are not formed, i.e., the items in the list do not serve as stimuli for succeeding responses but that, instead, responses are somehow learned to "self-generated sequential or spatial" symbols such as "first," "second," etc. Should such conclusions actually turn out to be valid, a great deal of traditional work on serial learning would have to be reevaluated, and much theoretical speculation would have been generated for naught. In view of the enormous amount of such literature and speculation, Slamecka's arguments should not be allowed to stand without a challenge.

It is my contention that Slamecka has concocted an ingenious and plausible, but incorrect, account of the process of serial learning. Much of what he has to say is indisputable, and, granting his assumptions, much of his argument appears unassailable. Closer examination, however, suggests that the attack is directed at a series of straw men, and that the traditional views Slamecka hopes to deny are hardly affected. It is, however, proper to point out that the fresh approach taken by Slamecka has produced a number of dividends in our understanding of serial learning. These will be duly noted.

In this commentary on Slamecka's analysis I shall first point out some general considerations and then comment on his specific arguments and supporting experimental evidence.

GENERAL CONSIDERATIONS

Slamecka argues that traditional psychology has made a mystery of the common finding that different parts of a list are learned with different efficiency, with the point just beyond the middle of a list being the most difficult to master. Theorists like Hull (1935) tried to explain such findings in terms of postulate systems and used data from experimental tests of derived theorems as support for the postulates. He regarded the differential difficulty demonstrated in serial list learning as a real and meaningful problem, a problem that called for solution along theoretical lines. Slamecka, in his approach, completely reverses the situation. He *assumes* the differential difficulty as a given that requires no explanation and proceeds to use this empirically established difficulty as the basis for his explanation of what goes on in serial learning. This appears to be a clear case of begging the question. Slamecka's raw empiricism solves the theoretical problem by ignoring it. The nature of association is reduced to a lawless cataloging operation.

The explanation offered by Slamecka amounts to a statement that when a learner masters a serial list he really learns the ordinal position of each item

as a kind of geographical locus or point in some kind of space. Each item is learned without regard to any possible associations it might have with preceding items or with those that follow. It would be fair to ask Slamecka to reply to such questions as: "What is the seventeenth letter of the alphabet?" or "Who was the twenty-first president?" without any attempt at such operations as reciting prior letters or taking advantage of any other known common associations like "Lincoln was 16; Johnson came after Lincoln, so he was 17; then came Grant, 18; then Hayes, 19; Hayes was followed by Garfield, 20, and, ah, yes, Garfield, Arthur, 21st."

The proposition that we learn serial lists of nonsense syllables through somehow noting their order without associating the syllables with each other is certainly challenging even though not readily acceptable. There is no question that the first and last syllables are learned relatively quickly and that they are recognized and probably labeled as "first" and "last" very early in the process. But to think that such essentially superfluous information is the essence of learning serial lists is to ignore the fact that in many serial tasks, for example, in the playing of musical compositions, reciting poetry, etc., the learner who forgets commonly goes back to some known point and starts afresh depending upon some kind of response-produced cueing process to help him continue.

When a learner does master a relatively short list, e.g., 10 items, he may very well be able to tell which syllable occupies what position, but not because he learned them according to the numbers. If we ask such a learner for the eighth syllable, he can quickly rehearse the list backwards, i.e., he "knows" the last syllable well and can probably generate its stimulus. That may enable him to think of the previous item. Similarly he can tell us the second and the third. With longer lists we could expect the process to break down just as we can expect our college-educated colleagues not to know the sixteenth letter of the alphabet or the twentieth unless they start from some

fixed point and count while associating one letter with its predecessor.

SPECIFIC ARGUMENTS

Slamecka attacks the principle of remote association on three fronts, those of "revised lists," association tests, and intrusion errors. Each of these has been used in the tradition to support or justify the concept of remote association. However, even confirmed believers in such associations have admitted that the first two sources of evidence are not strong and merely support the concept. Slamecka, though recognizing such admissions, mounts his strongest attack on the first two lines of evidence. I shall examine each argument in order.

Method of Derived Lists

Slamecka argues that studies using "derived" lists do not provide evidence for RAs; that, instead, they merely prove that a subject could instruct himself to take advantage of known facts about the construction of the list, e.g., a subject learning a first-degree remote derived list might simply say to himself: "Every second one." Slamecka goes on to argue that learning derived lists is no more efficient than learning control lists. Actually these arguments are beside the point as the evidence, since Ebbinghaus, has shown that derived lists may inhibit performance as well as facilitate it. There is no way of telling whether a derived list will be easy or difficult to learn in any given instance. All that Ebbinghaus claimed was that a first-order derived list would be learned better than a second- or third-order list. Slamecka reports an experiment in which he compares a derived list with a control list and finds no differences. Such a finding is simply irrelevant.

In designing his study Slamecka created what he calls a "modified" or "average" derived list which consists of three kinds of items—some are direct associations from the original list, some are one degree remote, others are two degrees remote. This he argues prevents the subject from using some easy formula

approach such as "skip every other one." Certainly such mediational devices should be controlled, but to confuse the subject with an arrangement where three different kinds of mediators might well occur to the subject is certainly not the proper procedure for getting rid of mediators. To make matters worse, Slamecka uses letters of the alphabet which already have strong direct associations, which confuse the issue further.

Association Method

Here Slamecka argues that serial position effects are the real explanation of so-called RAs. Syllables are learned best at both ends of a list and therefore have a higher availability if called for in association tests. How such an argument explains RAs is difficult to understand as subjects should not know the middle items well and would be unable to respond with them to early items. Only the first few items should have any RAs at all, and these should be from the other end of the list instead of from the middle. In practice, most RAs come from the middle of a list.

To test his argument, Slamecka designed a procedure of random exposures of six nonsense syllables at different frequencies and then tested for associations. His procedure called for 25 exposures of one syllable, 10 exposures of two others, 3 of two more, and only 1 of the sixth syllable. To take this arrangement of frequencies seriously as an analogue of serial learning, we would have to assume that subjects normally take 25 trials to learn a six-syllable list and that the first syllable is learned 25 times better (faster) than the middle one. Such assumptions strain credibility, but apparently are required to produce the data Slamecka then claims demonstrate that serial position effects account for RAs and not vice versa. The procedure also results in a finding of almost as many direct backward associations as forward associations (45-40) and more backward remote associations of first, second, and third degree than forward associations—a result yet to be reported in any learning literature. In

fact, Slamecka finds more backward than forward associations if we compare the four degrees of remoteness possible for backward associations with the first four degrees of forward remoteness. This fantastic result belies any comparability of such artificial procedures and normal learning. The argument from this analogy appears to be strikingly irrelevant to the issue. It should be noted that in this attack Slamecka appears to be arguing that RAs are given by the subjects because the responses are known and available, i.e., RAs are a function of availability. In the next section we will note that Slamecka argues that RAs occur because the responses are *not* known—the responses that are known and identified by position are inhibited by the subjects because they are known to be incorrect. The argument appears heavily ad hoc.

Method of Anticipatory and Perseverative Errors

Under this heading Slamecka offers no new experiment; what he does instead is to report data from a previous experiment which demonstrate the traditional findings of RAs. Slamecka explains these RAs, however, as matters of inadequate identification of the position of different responses. To learn a list, runs the argument, subjects must learn not only the item itself, that is, become familiar with the item per se, but must also learn its position. Until an item is learned (i.e., becomes familiar) it cannot be ventured as an association, remote or otherwise. As learning proceeds with the first and last items becoming identified and eliminated as potential intruders, the field from which RAs can be drawn is narrowed. As each item becomes identified with a certain position, it is no longer ventured as an RA. This explains why the middle of the list provides most of these incorrect, "remote," associations or "intrusions." While the items are still unlearned, states Slamecka (1964),

At this early stage he is more likely to emit remote association of a higher degree than later, since his ignorance of the list is maximal and his guesses are less constrained.

From this remark I gather that the subject is unaffected by any serial position effects and that his guessing is just that the subject is just as likely to respond with any one item as any other. Of course, with succeeding trials the number of intrusions decreases, and eventually there are none. The traditional pattern of RAs is explained away by Slamecka as a simple artifact or by-product of the assumption that when the subjects learn a list they learn it from both ends at first and slowly close in on the middle items. As the middle range is narrowed the degree of remoteness available decreases; this automatically results in more first-order RAs, fewer second order, still fewer third order, and so on. While there is some merit to this argument, it is not at all a fact that a list is learned in the symmetrical way Slamecka requires. Commonly, there is a clear advantage to the early items over the late ones. In any event, on the very first trial, there can be no such cumulative squeeze play as Slamecka describes. If RAs are formed on such a first trial, and if they are from an unconstrained range, they should not follow any particular pattern of distribution; but, if Slamecka is right, they should occur more or less randomly.

The Slamecka argument suggests two testable deductions: (a) if the subjects are asked to learn a list of already familiar items, say common words, they will not waste time on a familiarization stage—they can become reasonably familiar with the items in one trial. All of the words in a short list should then be more or less available as intrusions. If only one trial is given, there can be no piling up of short-range RAs. Further, if one word, e.g., the third word of the list, is then presented to the subjects with a request that they name the word that followed it, they should be free to respond with any of the words of the list except the key word itself. If they respond randomly, then Slamecka can be credited with a correct analysis. If, on the other hand, the pattern of responses follows the traditional curve (see Bugelski, 1950), then his argument must be rejected. (b) Because the subjects will

undoubtedly learn some of the words and be able to report them correctly, i.e., in their proper locus in the list, they should not use such known responses as RAs. If a given subject, for example, knows the tenth word to be the tenth and not the fourth, he should not use it as a response to the third word. In short, the Slamecka analysis calls for a negative correlation between the frequency of correct responses (correct position location) and the frequency with which such words are reported as RAs.

EXPERIMENTAL REBUTTAL

To test these deductions the author selected a list of 10 common 4-letter adjectives from a collection prepared by Castaneda, Fahel, and Odom (1961). None of the 10 adjectives selected was an associate of any of the others as tested by Castaneda, Fahel, and Odom; but as an additional precaution, three arrangements of the 10 words were prepared to counteract possible tendencies of the subjects to form meaningful associations between the words because of some potential "belongingness."

Procedure

Three classes of introductory psychology students were used as subjects. In each class, the author, serving as experimenter, told the subjects that they were to participate in a typical serial learning experiment, that they would be shown a list of 10 common English words at 3 seconds per word, and that they were to learn the words in serial order. Nothing was said about the number of trials, and it can be presumed that most of the subjects expected several trials. The words were printed in 4-inch block letters and could be seen easily by all subjects. Each stimulus card was exposed for 3 seconds. As soon as the last card was shown, the experimenter asked the subjects to write down the word that followed "_____." In each class a different word order was used, and in each case the third word in the list was pronounced as the cue or stimulus word to be responded to. Nothing was said about its position. The subjects were instructed to write down the correct word if possible and if not to write down the most likely word they could think of from the list just shown.

TABLE 1
RAs TO THE THIRD WORD IN A
10-WORD SERIAL LIST

Word in list	Degree of remoteness	Number of responses
1	-2	2
2	-1	19
3	—	—
4	0	— ^a
5	1	77
6	2	28
7	3	19
8	4	16
9	5	5
10	6	11
Total		177

^a The fourth word would be a direct association. It was reported correctly by 178 of the 355 subjects.

As soon as the subjects had written their responses, the experimenter instructed the subjects to write down the 10 words in order and to leave blanks where they could not recall the proper word.

Results

The total number of subjects was 355, and the total number of responses for the first test response was also 355. Of these 178 were correct, and the remaining 177 were incorrect. As seen in Table 1 there were 21 backward associations and 156

forward associations distributed according to the traditional pattern of mostly first-order remoteness, fewer second, less of third, and so on. Six degrees of remoteness were possible, and all were represented. The presence of such a pattern after the first trial should adequately rebut Slamecka's expectation that the responses should be randomly distributed. The distribution is significantly different from a random assortment of responses ($\chi^2 = 21.49$, $df = 5$, $p = .005$).

To answer the proposition that the subjects in such a situation refrain from using items whose position has been learned, a Pearson r was calculated between the number of responses correctly identified in the second test (where the subjects tried to report the whole list) and the frequency with which such responses appeared as RAs. Slamecka would have to assume a negative r . The actual r was positive, .593, and significant at the .05 level of confidence. It might be noted that the last word, although learned better than the ninth word (see Table 2), appeared as an RA twice as often.

The results just reported indicate that RAs are formed even on the first trial with familiar material. There can be no question of shrinking ranges forcing out some artifactual array of RAs. The data

TABLE 2
CORRECT AND ERRONEOUS RESPONSES IN A RECALL TEST
AFTER ONE LEARNING TRIAL

Word in list	Reported correctly	Position correct	Correct response, wrong position	Incorrect response, wrong position	Not reported
1	348	0	0	5	2
2	267	4	2	40	46
3	241	27	10	94	10
4	164	14	26	82	83
5	96	25	19	138	102
6	63	19	25	113	154
7	46	24	24	98	187
8	38	27	17	113	187
9	45	30	15	77	218
10	76	53	3	78	198
Total	1384	223	141	838	1187

Note.— $N = 355$.

gathered also suggest that one should be careful about postulating the order of learning in terms of some pattern of beginning and end first, then the middle. It is true that there is some small advantage of the very end of the list over the ninth and eighth words, but the striking feature of the data is the preponderant mastery of the early part of the list. Such mastery of the first two words could account for their infrequent appearance as backward RAs, thus supporting Slamecka's position to some degree.

Certain interesting data were available from the records of the second test where the subjects tried to recall the entire list. These data, shown in Table 2, summarize what the subjects can learn in one trial and how various responses are distributed. Remembering that these data are for a total recall operation and no longer a simple test of the response to the third word, we can note the subjects learned almost half of the task in the one trial given. Most of the correct responses were made up of the first four items. On 223 occasions a word (see Column 3) was reported correctly even though the stimulus word was missing or incorrect. Slamecka might argue that this shows that learning serial material is learning by the numbers. While this interpretation cannot be denied, it is possible to explain the occurrence of any word at any correct position as a confluence of RA effects. In the fourth column are listed those cases in which a correct response was made to a proper stimulus although neither was in the correct position. These data might be taken to support the association doctrine Slamecka denied. The fifth column shows that each word in the list was reported in one place or another incorrectly by at least some subjects. It is surprising to note how frequently the last item was reported as other than last. This last item apparently was familiar enough to be mentioned, but had not been identified as "last." In the last column of Table 2 are listed the failures to use the word in any form at all. The prevailing trend shown is an inability to come up with the later words in the list.

GENERAL CONCLUSIONS

The experiment reported above appears to answer Slamecka's criticisms of the traditional doctrine of remote associations on every count. By using meaningful (familiar) words, it is possible to take advantage of first trial responses which are uncontaminated by the effects of repeated trials. The argument advanced by Slamecka relies heavily on cumulative trial effects as well as rapid learning of the end of the list. Neither of these factors can be responsible for the present results.

The review of Slamecka's arguments showed them to be less directly pertinent than a first reading might suggest. The experiment reported here throws additional doubt on Slamecka's hypothetical position curve and makes his arguments from repeated trials irrelevant. The simple demonstration of the traditional curve of RAs developed after one run through a serial list should keep RAs in the category of useful psychological constructs for a while longer. The doctrine of direct association of one word with another which Slamecka also challenged hardly needed support, but the data reported in the present study revealed enough instances of word-to-word association regardless of serial position to justify the confidence of former researchers in the reality of such associations also.

REFERENCES

- BUGELSKI, B. R. A remote association explanation of the relative difficulty of learning nonsense syllables in a serial list. *Journal of Experimental Psychology*, 1950, 40, 336-348.
- CASTANEDA, A., FAHEL, L. S., & ODOM, R. Associative characteristics of sixty-three adjectives and their relation to verbal paired-associate learning in children. *Child Development*, 1961, 32, 297-304.
- HULL, C. L. The conflicting psychologies of learning: A way out. *Psychological Review*, 1935, 42, 491-516.
- SLAMECKA, N. J. An inquiry into the doctrine of remote associations. *Psychological Review*, 1964, 71, 61-76.

(Received June 2, 1964)

**"The first sane, good,
worthwhile book written
in the field of homo-
sexuality. . . .**

*This book inspires hope for
the doomed," said a well-known
psychiatrist after reading*

TOWARD AN UNDERSTANDING OF HOMOSEXUALITY

**by Dr. Daniel Cappon,
M.B., F.R.C.P. (Edin.)**

Dr. Cappon, Professor of Psychiatry and chief of the Psychodynamics and Psychopathology Section at the University of Toronto, has a large private practice. In this landmark book Dr. Cappon contends that homosexuality is a symptom of faulty development and adaptation, that its causes are reversible and consequently it *can* be cured by psychotherapy. As a physician in psychological practice he has diagnosed and treated some 200 homosexual patients.

Dr. Cappon corrects many common misconceptions and provides deep insights into the moral, historical, sociological and religious as well as clinical aspects of homosexuality.

This honest, realistic and optimistic new book offers counsel to doctors, clergymen and others who are called upon to help people with homosexual problems. *Bibliography*, \$6.95

at your bookstore
or Dept. 301

PRENTICE-HALL, INC.

Englewood Cliffs, N. J.

THE BRITISH JOURNAL OF PSYCHOLOGY

Edited by BORIS SEMEONOFF

Vol. 55 Part 4 November, 1964 25s. net

OBITUARY NOTICE:

CHARLES WILFRED VALENTINE.

MICHAEL ARGYLE. Introjection: A form of social learning.

P. M. A. RABBITT. Ignoring irrelevant information.

G. S. TUNE. Sequential errors in a time-sharing task.

L. TEFT, S. WAPNER, H. WERNER and J. H. MCFARLAND. Relation between perceptual and conceptual operations: Numerical distance and visual extent.

R. CONRAD and A. J. HULL. Information, acoustic confusion and memory span.

M. HAMMERTON and A. N. TICKNER. Transfer of training between space-oriented and body-oriented control situations.

T. X. BARBER and D. S. CALVERLEY. Experimental studies in 'hypnotic' behaviour: suggested deafness evaluated by delayed auditory feedback.

T. M. CAINE and K. HOPE. Validation of the Maudsley personality inventory E scale.

F. I. M. CRAIK. An observed age difference in responses to a personality inventory.

R. N. HUGHES. Responses by the ferret to stimulus change.

R. W. PICKFORD. A deuteranomalous artist.

ANN D. M. DAVIES. Season of birth, intelligence and personality measures.

CYRIL BURT. Baudouin on Jung.

PUBLICATIONS RECENTLY RECEIVED.

OTHER PUBLICATIONS RECEIVED.

The subscription price per volume, payable in advance,
is 75s. net (post free).

Subscriptions may be sent to any bookseller or to the

CAMBRIDGE UNIVERSITY PRESS
Bentley House, Euston Road, London, N. W. 1

PSYCHOLOGICAL REVIEW

ATTRIBUTE- AND RULE-LEARNING ASPECTS OF CONCEPTUAL BEHAVIOR¹

ROBERT C. HAYGOOD²

AND

LYLE E. BOURNE, JR.

*University of Utah**University of Colorado*

Analysis of the structure of concepts discloses 2 major features: (a) the relevant attributes, and (b) the conceptual rule by which the attributes are combined to form the concept. 2 experiments were performed to separate experimentally attribute identification and rule learning under a variety of 2-dimensional conceptual rules such as conjunction and inclusive disjunction. The results indicated that the rules differ in difficulty initially regardless of whether or not the relevant attributes are known, but that the differences decrease across successive problems. Further, it was found that knowledge of the rule represents valuable information to S which improves performance significantly. The findings were interpreted as demonstrating the validity of the analysis of conceptual behavior into attribute- and rule-learning components, and suggestions for future research to explore rule learning were formulated.

Problems used in experimental studies of conceptual behavior have two major features, either or both of which may be initially unknown to the subject. First, there are the stimulus characteristics which make up the specific concept-to-be-learned. Conventionally, these are called the relevant attributes; in commonly used geometric designs, they are values on relevant stimulus dimensions. Second, there is the general form or type of concept, represented by a rule which combines or otherwise elaborates the relevant attributes to define the con-

cept. For example, a bidimensional principle for classification might specify that "all red square figures are examples of the concept." In such a concept color and form are the relevant dimensions, while red and square are the relevant attributes or values of the dimensions; the conceptual rule is conjunction or joint presence of attributes.

Attribute Identification

In most concept-learning studies, particularly the concept identification experiments (e.g., Archer, Bourne, & Brown, 1955; Bourne, 1957), interest has centered primarily on the discovery or identification of relevant attributes. Typically the general form of solution is described and illustrated for the subject with preliminary instructions and practice problems and, as

¹ This report is Publication 48 of the Institute of Behavioral Science, University of Colorado. The research was supported in part by Grants MH-01759-06 and MH-08315-01 from the National Institute of Mental Health, United States Public Health Service.

² Now at Kansas State University.

such, constitutes a "given condition." Indeed, the majority of these experiments have employed simple and familiar unidimensional or conjunctive concepts or solutions. Once the unknown relevant attribute(s) is (are) discovered, through an inductive process based on the subject's observations of a series of examples and nonexamples of the concept, the problem is to all intents and purposes solved.

Correspondingly, concept-identification studies have explored variables which appear to be important to the identification of relevant attributes, such as number of relevant and irrelevant stimulus dimensions (Walker & Bourne, 1961), amount of intra- and interdimensional variability (Battig & Bourne, 1961), and redundancy between dimensions (Bourne & Haygood, 1959; Haygood & Bourne, 1964). Theoretical interpretations of concept identification have been based largely upon processes involving stimulus features of a single problem, such as the conditioning and adaptation of relevant and irrelevant cues (Bourne & Restle, 1959) and the sampling and testing of hypotheses about these features (Restle, 1962; Trabasso & Bower, 1964).

Conceptual Rules

Aside from incidental comparisons of the difficulty of unidimensional and conjunctive concept identification (Bourne & Haygood, 1959; Walker & Bourne, 1961), there has been no concern in this series of studies with type of concept or conceptual rule as a variable. The tacit assumption has been that identifying relevant attributes will be affected in much the same way by important variables regardless of the rule. Indeed, preliminary evidence, showing that the effects of number of relevant and irrelevant dimensions are the same for concepts

based on conjunctive and biconditional rules, has been reported (Kempers & Bourne, 1963). There is, however, some recent evidence (e.g., Conant & Trabasso, 1964; Hunt & Kreuter, 1962; Neisser & Weene, 1962; Shepard, Hovland, & Jenkins, 1961) to indicate that, under certain circumstances, types of concepts differ considerably in their difficulty. For this reason, it seems important to consider in detail the role that conceptual rules may play in concept learning. Although there are many rules for combining stimulus attributes and generating concepts, this paper will be limited to a system of constructing nominal, or discrete, concepts with, at most, two relevant attributes.

Consider first a closed stimulus population (Hovland, 1952) generated by x independent dimensions of y values each, consisting of the y^x possible combinations of stimulus attributes. When both x and y are two or more, it is possible to select two values from separate dimensions to be focal attributes. This selection maps the entire stimulus population onto the four categories or contingencies defined by the presence and absence of the focal attributes. For example, if redness (R) and squareness (S) are chosen as focal attributes, the four contingencies so defined are RS , $R\bar{S}$, $\bar{R}S$, and $\bar{R}\bar{S}$.³

When these two focal attributes are further selected to be relevant attributes in defining a concept, the four contingencies are mapped onto a two-response system: examples and non-

³ The bar over a symbol stands for "not," hence the categories read "red square"; "red, not square"; etc. In logical terminology the symbols T (true) and F (false) are used, representing the presence and absence, respectively, of the attribute specified. Thus in general terms, the contingencies are TT , TF , FT , and FF .

TABLE 1
FOURTEEN MAPPINGS OF A STIMULUS POPULATION WITH TWO FOCAL ATTRIBUTES
ONTO A BINARY RESPONSE SYSTEM

Dimensions and levels		Pattern sets			Partitions													
Color	Form																	
R	S	Contingencies	Two level	Three level	A	B	C	D	E	F	G	H	I	J	K	L	M	N
T	T	RS	RS	RS	+	+	+	-	+	+	+	-	-	-	+	-	-	-
T	F	R \bar{S}	RT	RT, RC	+	+	-	+	+	-	-	-	+	+	-	+	-	-
F	T	$\bar{R}S$	GS	GS, BS	+	-	+	+	-	-	+	+	-	+	-	-	+	-
F	F	R \bar{S}	GT	GT, GC, BT, BC	-	+	+	+	-	+	-	+	+	-	-	-	-	+

Note.—The following abbreviations are used: S—square, T—triangle, C—circle, R—red, G—green, and B—blue.

examples of the concept (positive and negative response categories, respectively). This second mapping creates 16 binary partitions of the stimulus population, two of which are trivial because they place the entire population in either the positive or negative response category. The remaining 14 nontrivial partitions are defined by the unique distribution of the four attribute contingencies (RS, R \bar{S} , $\bar{R}S$, and $\bar{R}\bar{S}$) into either the positive or negative category. The partitions are shown in Table 1. Capital letters are assigned arbitrarily to the 14 mappings of the stimuli onto the response system, with "+" indicating those patterns which are positive instances and "-" indicating those which are negative.

The number of stimulus patterns which are instances of each contingency is, obviously, a function of both interdimensional (x) and intradimensional (y) variability. Manipulation of interdimensional variability does not affect the proportions of patterns within each contingency, whereas a change in intradimensional variability clearly does. The pattern sets resulting in the cases of two-value and three-value dimensions are presented in Table 1.

Neisser and Weene (1962) have shown that there are only 10 different rules within this set of 14 mappings. This results from the fact that the following pairs are precisely the same except for a change of relevant attributes: B and C, E and G, H and I, and L and N. The 10 remaining mappings fall into 5 complementary pairs having the property that any instance which is positive under one member of the pair is a negative instance under the other. Table 2 describes these complementary pairs both symbolically and verbally.

Levels of Rules

A symbolic description, using the operations of negation, conjunction, and disjunction as primitives, suggests that the 10 mappings may be further broken down into levels based on the complexity (or length) of the descriptive expression. The unidimensional (Level I) represents the simplest, followed by those six expressions which involve single conjunctions or disjunctions of both relevant attributes (Level II). Finally, the most complex expressions entail both conjunctive and disjunctive operations (Level III). Neisser and Weene pointed out that the successive levels form a

TABLE 2
CONCEPTUAL RULES DESCRIBING PARTITIONS OF A POPULATION
WITH TWO FOCAL ATTRIBUTES

Parti- tion	Basic rule			Parti- tion	Complementary rule		
	Name	Symbolic description ^a	Verbal description		Name	Symbolic description ^a	Verbal description
E	Affirmation	R	All red pat- terns are ex- amples of the concept.	H	Negation	\bar{R}	All patterns which are <i>not</i> red are ex- amples of the concept.
K	Conjunction	$R \cap S$	All red <i>and</i> square pat- terns are ex- amples.	D	Alternative denial	$R \mid S$ $[\bar{R} \cup \bar{S}]$	All patterns which are <i>either not</i> red <i>or not</i> square are examples.
A	Inclusive disjunction	$R \cup S$	All patterns which are red <i>or</i> square <i>or</i> <i>both</i> are ex- amples.	N	Joint denial	$R \mid S$ $[\bar{R} \cap \bar{S}]$	All patterns which are <i>neither</i> red <i>nor</i> square are ex- amples.
C	Conditional	$R \rightarrow S$ $[\bar{R} \cup S]$	If a pattern is red <i>then</i> it must be square to be an example.	L	Exclusion ^b	$R \cap \bar{S}$	All patterns which are red <i>and not</i> square are examples.
F	Biconditional	$R \leftrightarrow S$ $[(R \cap S) \cup (\bar{R} \cap \bar{S})]$	Red patterns are examples <i>if and only if</i> they are square.	J	Exclusive disjunction	$R \oplus S$ $[(R \cap \bar{S}) \cup (\bar{R} \cap S)]$	All patterns which are red <i>or</i> square <i>but</i> <i>not both</i> are ex- amples.

^a Symbolic descriptions using only the three basic operators, \cap , \cup , and negation, are given in brackets.
^b There is no special symbol for exclusion in general use.

hierarchical structure, with the elements of Level II expressions being the Level I statements and the elements of Level III being Level II statements. Further they offer the hypothesis that the order of structural complexity within the system will be reflected in an order of difficulty when actual concept problems are given to human subjects for solution.

It is clear that the ordering of concepts in this analysis arises from the selection of negation, conjunction, and disjunction as basic operations. Hunt (1962) has adopted a more elaborate system of operators to describe the same expressions or conceptual rules; in fact, by reference to textbooks in mathematical logic (e.g., Quine, 1958), special operators or symbols for all but one of the rules of Table 2 can be found. There is nothing in Hunt's

system, other than the difference between concepts based on one as opposed to two relevant attributes, which would suggest any ordering of complexity. Cases in which these operators differ from the expressions provided by Neisser and Weene can be seen in Table 2; Level III concepts can be recast within the more elaborate system so as to *appear* of the same order of complexity as Level II concepts.

Neisser and Weene, however, report a study of the relative difficulty of attaining concepts at the three levels implied by their system, and the results clearly support the hypothesis that concepts of hierarchically higher levels are more difficult to attain than those of lower levels. In addition, although there is a general improvement in performance across two suc-

cessive concept problems of the same type, the difference in difficulty is still evident on the second problem. Exploring fewer different concept types, Hunt and Kreuter (1962) found similar results.

Rule Learning

The analysis above immediately raises the question of why, from a psychological point of view, higher level concepts are more difficult to attain. Hunt offers an interpretation based on the fact that any deterministic rule can be expressed as a tree of decisions. The tree specific to any rule is learned by the subject, assumedly through the utilization of a general strategy applicable to all conceptual problems. Theoretically, the subject begins by trying to solve the problem as a conjunction of attributes, but, in doing so, may uncover (when a nonconjunctive solution holds) subproblems which require special treatment. The general strategy may be realized in a computer program which yields results close to the order of concept difficulty observed in real data (Hunt & Kreuter, 1962). Neisser and Weene, on the other hand, lean toward an interpretation based on the hierarchical organization of concepts, which pictures the subject as working up from simple to more complex groupings of stimuli. To identify correctly an instance of a Level III concept, the subject must have Level II concepts available as components. Similarly, to utilize Level II concepts, the subject must be familiar with Level I components. Thus to attain a complex concept, the subject must use and therefore have attained concepts at a lower level. This interpretation implies that the subjects do not learn Level III concepts as such, but rather construct or induce them from their component parts.

There are at least two other con-

siderations relevant to the available data, neither of which is necessarily incompatible with the aforementioned interpretations of the subject's behavior. The first is suggested by an informational analysis of problems, which provides one means of measuring the complexity of decision trees corresponding to different rules. Level I concepts have one relevant dimension, whereas Level II and III concepts have two; number of relevant dimensions (or bits of relevant information) has been shown to be a powerful determiner of task difficulty (e.g., Bulgarella & Archer, 1962). The greater difficulty of Level III as compared to Level II concepts may result from differences in the mapping of the four stimulus contingencies as shown in Table 1 (i.e., TT, TF, FT, and FF) onto the two response categories. All Level II concepts involve a 3:1 split of the contingencies, whereas Level III concepts are based on a 2:2 split. In a sense, there is more uncertainty in the stimulus-category system of Level III concepts (Garner, 1962). This uncertainty is reflected in the lack of homogeneity or commonality among individual stimulus patterns in either the positive or negative category of Level III concepts. Indeed, each category within Level III concept types consists of two subsets whose members have, by definition, no relevant attributes in common. Thus the highly efficient strategies based on discovery of common attributes (Bruner, Goodnow, & Austin, 1956) are eliminated, and must be replaced by different, and perhaps more difficult and less efficient, strategies based on multiple contingencies. This interpretation suggests that differences in difficulty among concept levels (and perhaps within levels as well) result from differences in stimulus uncertainty—in effect, from

differing information requirements to solve the problems.

It is also possible that the observed ordering between rules is little more than a reflection of differing familiarity with the rules and their appropriate strategies. With equal familiarity, concept forms at Levels II and III may be equally difficult to attain. Suggestive of this is Wells' (1963) demonstration that the evident preference of naive subjects for conjunctive as opposed to disjunctive concepts in problems which could be solved by either rule (Hunt & Hovland, 1960) can be modified by preliminary training and familiarization with disjunctive concepts. Extensive training on several mappings of Table 1 may produce equal facility with each type of concept, thus modifying the hierarchy described by Neisser and Weene (1962).

Two Aspects of Conceptual Behavior

With the exception of that of Conant and Trabasso (1964), each of the studies of comparative rule difficulty discussed above has required the subject both (a) to identify the relevant attributes, and (b) to learn the assignment of levels on relevant dimensions to response categories as prescribed by the rule. That is, for the subject each problem had two unknowns. This confounding of attribute identification with rule learning creates certain difficulties in the interpretation of results. For example, it has not been established that the rules differ in difficulty in the absence of the necessity to identify relevant attributes. Clearly, the strategies appropriate for identifying relevant attributes differ among rules, and it may be that those which are useful in a conjunctive problem are easier to learn or employ than those associated with a biconditional. This possibility is supported by the finding of Conant and Trabasso that naive subjects learn to use a

positive-focus strategy in a conjunctive problem earlier than they learn to use a negative focus in a disjunctive problem. However, it is possible that attribute identification under different conceptual rules might not differ if adequate instructions and training on the relevant rule were given.

The foregoing discussion suggests that, paralleling the analysis of conceptual problems into rules and attributes, there is a distinction between two independent kinds of behavior: (a) rule learning, in which the subject attains or discovers the principle for partitioning the stimuli in a particular problem and acquires the rule in general form so that he can use it in any problem; and (b) attribute identification, in which the subject attains or discovers the relevant attributes. One major purpose of this paper is to provide some experimental evidence on the validity of this analysis, to demonstrate the separability of the two forms of behavior, and to indicate the unique features of each.

EXPERIMENT I

In the first experiment, the subjects were presented with a series of concept problems, each based on the same rule but with a different pair of relevant attributes. Each subject served in one of three main conditions which are labeled rule learning (RL), attribute identification (AI), and complete learning (CL). In the RL condition, the subject was given the names of the two relevant attributes at the start of each problem. His task was to learn the proper assignment of a set of stimulus patterns to the response classes according to some unknown but preselected rule combining these attributes. In the AI case, one rule was explained and illustrated for the subject at the outset, leaving him the task of identifying the attributes relevant to the concept in each problem. This case represents in essence the

experimental procedure used in earlier studies of concept identification. In the CL case, the subject was told neither the rule nor the attributes at the outset, but had to attack each problem given only a description of the stimulus population and the size of the concept (two relevant attributes).

One specific purpose for separating out RL was to determine whether the rules are intrinsically different in difficulty when the subject is not required to identify relevant attributes. If so, then it can be expected that, at least on the initial problems, different rules will yield different numbers of errors and trials to solution. In contrast, the AI condition (which provides thorough preliminary instruction on the rules) was designed to assess the relative difficulty of identifying relevant attributes under different rules. Whereas expectations regarding RL are not clear, there is some evidence in the literature to support the expectation that AI problems will differ in difficulty even though rule instruction is given. Bruner, Goodnow, and Austin (1956) have demonstrated the utility of a powerful holist strategy for identifying the relevant attributes of a conjunction. This strategy allows the subject to form and test a composite hypothesis which embodies on any trial all tenable solutions. Thus, with this strategy there is no need for the subject to consider independently all possible pairs of relevant attributes. Such a strategy is, in general, not appropriate to the disjunction or to any of the other rules to be studied in this experiment. The strategies which are available for these other rules provide no convenient way of forming a composite hypothesis which simultaneously represents all possible solutions.

The CL condition was included for two reasons. First, it provides continuity with the experiments of Neis-

ser and Weene (1962), and of Hunt and Kreuter (1962). Second, CL provides a base line for comparison to AI and RL to determine the effectiveness of instructions that, in one case, provide the relevant attributes and, in the other, the relevant rule. If knowledge of either feature is important to concept learning, performance on even the first problem should show significant differences between CL and the other two conditions.

To determine the effects of practice, a series of five problems having the same solution form was given in all conditions. Assuming the effectiveness of instructions, improvement across successive problems could be attributed to the following factors. In RL, to an increasing familiarity with the way a rule assigns stimulus patterns to the response categories; in AI, to the development and utilization of strategies appropriate to the identification of relevant attributes under a given rule; and, in CL, to both factors.

Method

Subjects and Design. Sixty students from introductory psychology classes participated individually as subjects and were assigned in order of appearance to one of 12 treatment combinations. All subjects were naive in that they had not previously participated in a concept-learning experiment.

The experimental design was a 4×3 repeated-measures factorial, incorporating concept problems generated by four different rules and three instructional conditions. The rules used, viz., conjunction, inclusive disjunction, joint denial, and conditional, were selected arbitrarily to include (a) only Level II types (Neisser & Weene, 1962) and (b) one pair of complements. The three instructional conditions were AI: the rule and all stimulus dimensions were described and illustrated, but the relevant attributes were unknown to the subject; RL: all the stimulus dimensions and the relevant attributes, stated in TT form, e.g., red square, were given, but the rule was not explained; and CL: all dimensions, but neither the rule nor the relevant attributes were described.

Each subject solved five successive problems utilizing the same rule within the same instructional condition. A different pair of relevant attributes was used in each problem, and five different sequences of pairs were used to avoid confounding with ordinal position.

Materials and Apparatus. The stimulus patterns were geometric designs, prepared on 4 × 6 inch cards and varying along four three-value dimensions. Dimensions and their values were: number (one, two, and three figures), size (large, medium, and small), form (square, triangle, and hexagon), and color (red, yellow, and blue). From Table 1 it is clear that the numbers of positive and negative instances within this population of 81 stimulus patterns vary markedly with the rule. In studies using only a single rule, the proportion of positive instances has been shown to be a significant determiner of rate of learning, even when the stimulus sequences have been equated for information content (Hovland & Weiss, 1953). To avoid confounding rule differences with this factor, extra cards were inserted in the basic deck when necessary to insure roughly equal numbers of positive and negative instances.

A sample card showing all possible values of the dimensions was left exposed for the subject's reference at all times. Special cards showing Venn diagrams of the conceptual rules to be used in the experiment (rule cards), and cards listing the pairs of relevant attributes to be used (attribute cards) were prepared. The subject sat at an ordinary desk facing the experimenter. Presentation of stimuli and recording of responses were performed manually.

Task and Procedure. The subject was required to sort, or classify, a series of visually presented stimulus patterns into two categories. The correct classification for any stimulus was determined both by the pair of attributes relevant in the particular problem and by the conceptual rule which specified the relationship between these attributes. In each problem two dimensions were irrelevant to solution.

At the outset, all subjects were given detailed oral instructions describing the stimulus population and the task. They were told that stimulus patterns would be presented one at a time and were to be sorted into two categories, viz., those that were and those that were not examples of the concept. Emphasis was placed on finding the two attributes which characterized the class of

positive instances; despite this emphasis on the positive category, the subjects were warned that both the presence and absence of particular attributes might be important. Instructions given to all groups were identical except for those portions restricted by the nature of the experimental conditions. The subjects were told that there would be five problems, each utilizing the same rule but each with a different pair of relevant attributes. Prior to the first problem, the subjects in the AI condition were given the appropriate rule card for their problem series. Before each problem, the subjects in the RL condition were given an attribute card listing the pair of relevant attributes. The attributes were read to the subject in the order in which they appeared on the attribute card. However, no emphasis was placed on the fact that one came first, or that the order in which the attributes appeared might be related to problem solution. The subjects in the CL condition were told only that two attributes would be relevant and that these would be combined by some kind of rule.

To each successive stimulus card, the subject was required to respond "yes" or "no" thus indicating whether it was an example of the concept or not. Following the subject's response, the experimenter provided feedback by saying "right" or "wrong" and by placing the stimulus card face up on the desk in an area marked YES or NO as appropriate. All problems were self-paced in that the subject was allowed as much time as desired to respond, and instructions stressed accuracy rather than speed.

The criterion of problem solution was 16 consecutive correct responses. When each problem was completed, the subject was reminded that the next problem would involve the same rule but a new pair of relevant attributes, and those in the RL condition were given a new attribute card. Programming of card sequences was randomized by shuffling the deck several times before each problem.

Results and Discussion

The mean numbers of errors and trials to solution observed for all five problems in each of the 12 main conditions of the experiment are reported in Table 3. Overall analysis of variance on errors demonstrated four statistically reliable effects. First, for all con-

cept types, CL, wherein both the rule and the relevant attributes were unknown, was more difficult than AI, which in turn yielded a greater number of errors than RL ($F = 19.34$, $df = 2/48$, $p < .01$). Second, there was steady improvement in performance over problems ($F = 13.15$, $df = 4/152$, $p < .01$) and no evidence of interaction between problems and instructional conditions. Third, rules differed generally in difficulty ($F = 12.24$, $df = 3/48$, $p < .01$) and according to the order observed by Hunt and Kreuter (1962). Finally, the interaction of rules and successive problems was reliable ($F = 2.91$, $df = 12/152$, $p < .01$) reflecting the fact that the difficulty of joint denial, disjunction, and conditional concepts changed markedly in comparison to conjunctive concepts. Identical results held for the analysis of trials to solution.

Further inspection and analyses of the data in Table 3 revealed that, in RL conditions, essentially perfect performance was reached by all subjects

on the fourth problem on all rules except the conditional. Initial expectation of and familiarity with the conjunctive rule was clearly evidenced by the fact that only one subject made any errors on these RL problems, and he made these only on two of the five problems. Fractional mean errors on later disjunction and joint denial problems resulted, in both cases, from a single subject making one or two errors per problem. Complete mastery of these three rules is demonstrated by the immediate solution of each problem by the subjects upon being provided with the relevant attributes.

The overall greater difficulty of the conditional rule appeared to result from two sources. First, the size and nonhomogeneity of the class of positive instances: inspection of Table 1 shows that the conditional rule classifies a greater number of instances as positive than any other and moreover uniquely requires both TT and FF (e.g., red square and not red, not square) to be grouped together. One implication of

TABLE 3
MEAN NUMBER OF ERRORS (E) AND TRIALS (T) TO SOLUTION IN EXPERIMENT 1

Rule	Instructional condition	Problem									
		1		2		3		4		5	
		E	T	E	T	E	T	E	T	E	T
Conjunction	AI	4.0	13.4	3.4	12.0	1.6	11.6	3.4	15.5	2.2	7.4
	RL	.2	.2	.4	.4	0	0	0	0	0	0
	CL	3.6	14.0	3.0	6.4	3.2	11.8	4.0	16.2	2.0	7.6
Disjunction	AI	12.1	32.8	7.6	32.0	4.0	18.0	4.2	9.8	4.2	16.0
	RL	8.0	33.3	.6	4.2	.2	1.0	.4	.4	.2	2.6
	CL	37.0	120.8	17.0	52.4	22.2	75.6	11.0	38.8	13.0	49.2
Joint denial	AI	9.2	32.6	11.4	35.6	6.4	21.5	8.6	27.0	3.4	16.6
	RL	8.6	28.6	3.4	11.8	1.6	4.6	0	0	.4	4.0
	CL	17.0	67.0	12.0	49.4	16.2	62.6	18.6	67.0	17.2	65.6
Conditional	AI	26.8	49.4	12.2	47.8	12.6	54.0	13.8	54.5	9.0	34.0
	RL	16.4	57.5	10.4	45.8	6.8	23.5	7.6	23.0	6.4	30.5
	CL	38.4	116.3	11.4	27.8	11.2	30.6	20.6	60.6	11.0	37.6

this is that focusing on the negative category (Hunt, 1962), which was not encouraged by the instructions of this experiment, would facilitate the learning of conditional concepts. Second, the asymmetry of the rule: whereas the partition generated by any of the other three rules is invariant with commutation of attributes, e.g., red and/or square is identical to square and/or red, only one of the two possible attribute orderings is correct for a conditional problem. The effect is to place special emphasis on the difference between TF and FT instances. Because the order of attributes was not stressed in the instructions, the subject was posed the additional task of making this basic distinction. Additional evidence on the difficulty of this task is provided by Experiment II.

Once the subject has learned the rule, it is obviously possible to categorize patterns without error in any RL condition. In contrast, AI requires, on logical grounds, the presentation of a certain minimal number of instances since the subject must discover which dimensions and values are relevant. Even the most efficient search strategy will be accompanied by a number of errors commensurate with the number of trials necessary to eliminate all but the correct pairing of attributes.

Although the specific patterns shown to each subject differed and were not recorded, it is possible to compute the expected minimal number of trials required for problem solution under the AI condition for any rule. These minima have been computed for the stimulus population used in the present experiment under the following assumptions: (a) the subject has limitless memory, i.e., never forgets information previously presented, and (b) the subject tests possible hypotheses using a strategy equivalent to the "simultaneous scanning" technique

(Bruner et al., 1956). The computed values are 4.7, 10.8, 9.7, and 8.0 for conjunction, inclusive disjunction, joint denial, and conditional, respectively. Inspection of Table 3 shows that this theoretical minimum is not attained for any rule, although it is approached for conjunctive problems. It is clear, both from the present data and from those of Bruner et al. that the subjects do not use simultaneous scanning. Any strategy adopted by the subject will require a greater number of trials than a maximally efficient scanning process, and the minimal number of trials will depend on the strategy actually used by the subject.

While it is impossible to determine exactly what strategies, if any, were used by the subjects, inspection of the error data suggests that the subjects were not responding in accord with randomly selected binary hypotheses prior to solution, for such a strategy would give an error total equal to about half the number of trials. Instead, the actual mean errors approaches the minimum commensurate with the *theoretical minimum trials* for all rules except conditional. In other words, it would appear that error performance in conjunction, inclusive disjunction, and joint denial was about as efficient at the end of five problems as it can be. The small ratio of errors to trials suggests that many subjects solve the problem in parts, for example, by initially discovering one but not both of the relevant attributes (Trabasso & Bower, 1964).

The overall difficulty of the conditional rule is reflected in performance within the AI condition. Numbers of trials and errors are greater over all five problems than for any other rule. The failure of the subjects to obtain conditional rule mastery in five consecutive RL problems suggests that part of the difficulty within the AI

condition results from a lack of understanding of the rule itself.

Except for conjunctive concepts, CL yields more errors and requires more trials than AI, even for the fifth problem. This result demonstrates the effectiveness of prior instructions about the rule, and provides further evidence that both RL and AI are involved in the solution of conceptual problems. The improvement in CL performance is consistent with the findings of Neisser and Weene (1962). It is indicative of the possibility that, with further training, the subjects will become as facile with problems of a non-conjunctive nature as they are with conjunctive problems at the outset—within the constraints of the amount of information required for problem solution. However, failure of the CL groups, in general, to attain the same level of performance as AI groups after several problems indicates that rule learning proceeds more slowly when the subject must, in the same problem, identify relevant stimulus attributes. Clearly, then, CL is more than the simple sum of AI and RL.

At least after the first problem, performance on the disjunctive and joint denial concepts is virtually identical, as would be suggested by the fact that they are complementaries. The difference between these concepts on the first problem, which is not statistically significant, may be attributable to the initial effect of instructions emphasizing the class of positive instances, but the evidence is far from compelling.

EXPERIMENT II

The RL condition of Experiment I provided training for the subject in the assignment of stimulus patterns (or attribute contingencies) to response categories by a single rule. If the subject can be trained to perform adequately on two or more rules, it

becomes feasible to present to him a conceptual task which may be labeled, for sake of consistency, rule identification (RI). In RI the subject is given the relevant attributes and is required to identify which of the known rules determines problem solution. The task, obviously, is designed as an analogue of AI problems, wherein the subject must identify which of several known attributes determines solution under a given rule. From Table 1 it is clear that the assignment of one or more of the four contingencies of relevant attributes will differentiate any two rules, and that presentation of all four contingencies will determine uniquely which rule is correct. Thus the subject's performance may be compared to an efficient algorithm, which solves such RI problems in four trials if each trial provides an example of a different contingency of the relevant attributes. In general, a certain minimal number of instances, positive or negative, is required to eliminate all but the correct solution for any RI problem, just as in AI. Assuming prior familiarization and equalization of rules through RL, however, no difference in rule difficulty would be expected.

The purpose of Experiment II is to investigate RI performance after prior RL training on four different rules selected to represent both Levels II and III from the Neisser and Weene (1962) hierarchy and a high degree of diversity in the way they assign stimulus patterns to categories.

Method

Subjects and Design. Twenty-four students from introductory psychology classes served as subjects and were assigned in order of appearance to six training groups. None of the subjects had participated in Experiment I; however, approximately three-fourths had participated in other concept-learning experiments, involving the

identification of attributes in unidimensional or conjunctive concepts. Because of task differences, and because these subjects were distributed without bias over experimental treatments, prior experience was assumed to have no effect on the present results other than possibly an overall improvement in the general level of performance.

The experiment consisted of two parts, an RL and an RI phase. In the RL phase, the subjects were presented with a series of concept problems based on three different rules, viz., disjunctive, conditional, and biconditional. The six possible sequences of rules were used, each for an equal number of subjects. Successive problems utilized different pairs of relevant attributes, which were named for the subject before each. The purpose of this phase was to train the subjects in the use of these three conceptual rules, plus a fourth, the conjunction, which was explained in detail and illustrated during initial instructions.

For the RI phase, a balanced incomplete-blocks design, with four rules, two pairs of relevant attributes, and two successive problems, was used. The two problems were based on two of the four rules on which the subject was trained during the RL phase. The two pairs of relevant attributes, which were named for the subject at the outset of each problem, were counterbalanced with problem order and with relevant rule.

Materials and Apparatus. The stimulus patterns were the same as those used in Experiment I. Sequences, however, were prepared on filmstrips, and the form of the subjects' responses changed from verbal statements to pressing buttons, labeled "yes" and "no," on a control panel. The apparatus was the same as that used in previous concept-identification studies (e.g., Bourne & Haygood, 1959). It was programed to present an immediate informative feedback signal of 1-second duration to the subject after each of his responses. The total time between the subject's response and the appearance of the next stimulus pattern was fixed at 5 seconds.

In preparing stripfilms, several constraints were imposed on the sequences of stimulus patterns, in addition to the equation of number of positive and negative instances. First, and most important, the regular occurrence of a representative of each of the four attribute contingencies shown in Table 1 was provided. Particular attention was paid to insuring that instances critical to the distinctions between the four rules occurred at least once during any sequence of 16 patterns (the

length of the criterion run). Second, runs of the same category (yes or no) were limited to five. Third, where patterns were repeated because of using extra cards to equate positive and negative instances, care was taken that no two consecutive patterns were identical. Fourth, each problem began with a pattern containing both relevant attributes, i.e., the TT attribute contingency.

Task and Procedure. The task and instructions were essentially those used in the RL condition of Experiment I, with the exceptions that (a) equal emphasis was placed on observing instances of both the positive and negative categories and (b) an elaborate explanation of each rule was given after performance on its associated problem. Detailed instructions were given the subject concerning the stimulus dimensions to be employed in the problems, the operation of the apparatus, and the nature of concepts and conceptual problems. During the instructions, the conjunctive rule was explained and used as an example, and the subject was given the conjunction rule card. The subject was told that there would be three training problems, each employing a different rule, and that subsequently there would be two problems in which the solution could be any one of the four rules.

The stimulus patterns were projected one at a time on the screen before the subject. To categorize a pattern, the subject responded by pressing either the yes button or the no button on the control panel below the screen. The correct response was then indicated by the illumination of a signal lamp above the proper button. All problems were self-paced, the subject being allowed as much time as needed to make any response. The criterion of problem solution was 16 consecutive correct responses. Any subject failing to solve the first problem in 45 minutes was arbitrarily deemed a nonsolver and dismissed from the experiment. Two subjects were discarded under this criterion; one other subject was eliminated for being unable to solve the third problem and being unwilling to continue in the experiment. Three other subjects were eliminated for reasons unconnected with the experimental treatment, such as apparatus malfunction.

Following solution of the first problem, the subject was given the proper rule card and a thorough explanation and further example of the rule. When the three RL problems were completed, the subject was reminded of all four rules and that the next two problems would each utilize one of the four rules he had just learned.

Results and Discussion

Mean numbers of errors and trials to solution of all problems are given in Table 4. Analyses of variance on the two dependent response measures showed identical outcomes; therefore, only significance tests based on errors are reported. During the RL phase, sources of variance identified with rules ($F = 9.19$, $df = 2/40$), ordinal position of problems ($F = 9.15$, $df = 2/40$), and their interaction ($F = 6.08$, $df = 4/40$) were statistically significant ($p < .01$). The initial order of difficulty among rules is precisely the same as that observed by Neisser and Weene (1962) and by Hunt and Kreuter (1962); these differences, however, are attenuated on the latter two RL problems. Because, in this phase, only one working problem of each type was given to each subject, ordinal position and interaction effects are attributable to a general transfer among rules, particularly from other rules to the biconditional, rather than (as in Experiment I) a specific process of learning how to solve problems characterized by a single rule.

Number of trials to present at least

one stimulus pattern representative of each type of attribute contingency sets one logical limit on the minimum information necessary to categorize all patterns correctly. These values, computed directly from the stimulus sequences used, are presented in Table 4. A comparison of trials to solution with this base line indicates that, with the possible exception of later disjunctive problems, the subjects do not solve problems with maximal efficiency during the RL phase. This is hardly unexpected in view of the training routine employed and the relative unfamiliarity of these rules at the outset.

There are significant differences in difficulty among conceptual rules in the RI phase ($F = 6.39$, $df = 3/18$, $p < .01$) with biconditional and conditional problems producing a greater number of errors than the conjunctive and disjunctive. The reliability of these differences is enhanced by the limited variability in performance on conjunctive and disjunctive problems, wherein solutions were attained in a minimal number of trials in 75% and 83% of cases, respectively. Identification of the conditional and biconditional rules was accomplished in a

TABLE 4

MEAN NUMBER OF TRIALS (TS) AND ERRORS (E) TO SOLUTION AND NUMBER OF TRIALS NECESSARY TO PRESENT AT LEAST ONE REPRESENTATION OF EACH ATTRIBUTE COMBINATION (TP) IN ALL PROBLEMS, EXPERIMENT II

Rule	Problem														
	Rule learning									Rule identification					
	1			2			3			4			5		
	E	TS	TP	E	TS	TP	E	TS	TP	E	TS	TP	E	TS	TP
Conjunction	—	—	—	—	—	—	—	—	—	1.8	8.0	8	3.2	5.8	5
Disjunction	6.3	19.5	7	2.3	6.9	7	2.8	8.4	5	1.5	4.2	5	2.3	7.2	5
Conditional	8.8	24.4	9	10.5	27.0	5	6.3	15.5	9	3.5	8.8	7	6.8	16.3	6
Biconditional	29.1	65.5	7	4.4	16.0	7	8.6	27.4	5	8.5	30.2	9	6.2	27.8	5

Note.—There were no conjunctive problems given during the RL phase.

TABLE 5

MEAN NUMBER OF TOTAL PRESENTATIONS (TP), PRESENTATIONS TO LAST ERROR (NP), ERRORS (E), AND RATIO OF ERRORS TO PRESENTATIONS (E/NP) FOR FOUR ATTRIBUTE COMBINATIONS, EXPERIMENT II

Attribute combination	Disjunction				Conditional				Biconditional				Conjunction			
	TP	NP	E	E/NP	TP	NP	E	E/NP	TP	NP	E	E/NP	TP	NP	E	E/NP
Rule learning, Problems 1-3																
TT	1.5	.4	.3	.80	3.0	.9	.6	.63	4.3	2.8	1.0	.36				
TF	2.3	1.4	1.2	.88	9.2	5.2	2.6	.53	8.0	6.4	3.5	.54				
FT	3.3	2.0	1.3	.65	3.1	1.3	1.1	.87	8.3	5.9	3.3	.57				
FF	4.9	2.5	.9	.36	6.6	6.3	4.1	.65	15.8	11.3	6.3	.56				
Rule identification, Problems 4-5																
TT	1.0	0	0	—	1.4	.3	.3	1.00	3.7	.7	.3	.43	3.0	.6	.4	.67
TF	1.0	.9	.5	.56	6.0	4.0	1.8	.45	5.7	5.2	2.4	.46	1.3	1.2	.9	.75
FT	1.4	1.3	.8	.61	1.5	1.2	.9	.75	7.9	3.8	2.1	.55	1.1	.9	.6	.67
FF	2.3	1.7	.6	.35	3.7	3.4	2.2	.65	11.7	5.0	2.5	.50	1.5	1.5	.6	.40

minimal number of trials by only 33% and 25% of the subjects, respectively.

As in Experiment I, errors were made on fewer than 50% of the trials, as is shown in Table 4. To determine the correctness of an assumption that this outcome resulted from learning the assignment to response categories of certain contingencies more rapidly than others, problem-by-problem analyses were made of (a) the number of presentations of each type of attribute contingency prior to the first trial of the criterion run, (b) the number of presentations of each type of attribute contingency prior to the last error on an instance of *that type*, and (c) the number of errors made on each type of attribute contingency. The results of those computations are shown in Table 5 for the three RL and the two RI problems separately. Although there are variations across problems and across attribute contingencies within problems, the proportion of errors on trials up to and including the last error trial computed separately for

each contingency (the ratio of *c* to *b* above) averages .56 in RL problems and .53 in RI problems. Thus, patterns apparently were assigned to categories with chance probability of being correct until the assignment was learned, after which no further errors were made.

Variation among attribute contingencies in number of trials to last error reveals some sources of rule difficulty. The scores identified with the TT combination may be discounted because (a) an instance of this type was the first presented in every problem, a fact commonly discovered and utilized by the subjects, and (b) the instance was positive for all rules, thus not one which distinguished among them. Differences among the three remaining contingencies were not great for disjunctive problems in the RL phase. On the average, the subject required presentation of about twice as many of each type as was necessary on logical grounds. On conditional rule problems, however, TF and FF required a greater number of

presentations to last error than did FT ($F = 3.71$, $df = 2/42$, $p < .01$). Both instances are critical to the disjunction-conditional distinction, their assignment to categories being reversed under the two rules; and the difference may be a reflection of the greater ease of learning or familiarity with disjunctions. Note also that any strategy which emphasizes the search for common attributes among instances of the positive class encounters difficulty with the conditional, which assigns patterns of the TT and FF types to the positive category. This is apparent in performance on biconditional problems, also, wherein the largest number of presentations to final error was associated with FF instances ($F = 3.07$, $df = 2/42$, $p < .10$) which, as in conditional problems, are assigned along with TT instances to the positive class. In this rule, of course, any instance representing one but not the other relevant attribute (TF, FT) must be categorized as a negative instance. Learning the assignment of these instances is apparently more difficult under the biconditional than under any other rule ($F = 3.25$, $df = 2/110$, $p < .05$).

A similar analysis of RI problems is presented in Table 5. Except for overall better performance, the attribute contingencies are aligned in order of difficulty much as they were in RL problems. Mean number of presentations is close to 1.0, and the error rate is approximately .5 for all contingencies within the conjunctive and disjunctive rules. Obviously these rules were completely mastered either during RL or through preexperimental experiences of the subjects. Only 5 of 24 subjects used more than the minimal number of trials required by the algorithm. The evidence strongly suggests that the subjects have acquired (a) the technique of reducing the population of stimulus pat-

terns into attribute contingencies and (b) the mapping of contingencies onto the response system associated with both the conjunctive and disjunctive rules.

Three to four presentations of TF and FF instances for the conditional rule and four to five presentations of TF, FT, and FF instances for the biconditional were required on the average before these RI problems were solved and the rule satisfactorily identified. Incomplete mastery of these rules appears to be the most prosaic explanation of poorer performance. While there is strong evidence for improvement on these rules, only 4 of 12 subjects for the conditional and 3 of 12 subjects for the biconditional solved in minimal trials. The fact that 6 of the 7 subjects performing optimally on conditional and biconditional problems solved their other RI problem in minimal trials suggests that they have learned to use, at least intuitively, the principle for collapsing the stimulus population to attribute contingencies and that the difficulty of conditional and biconditional problems stems primarily from a failure to learn completely the mapping of contingencies to responses necessitated by these rules.

GENERAL DISCUSSION

The preliminary analysis outlined in this paper discloses two aspects of conceptual problems as they are commonly posed for subjects in the laboratory, viz., the learning of rules and attributes. Research to date has been concerned largely either with (a) the processes by which subjects, given the correct rule, discover or identify relevant attributes of a concept, or (b) performance in situations wherein both the relevant attributes and the rule are unknowns. The present experiments considered the third case, attributes given and rule unknown, for purposes of logical and

empirical comparison with the first two. Primarily, this study provided a comparison of performance characteristics in RL and AI problems; its aim was to determine the advantages to be gained, if any, from a separation of attribute and rule aspects in studies of conceptual behavior. Secondly, interest centered on the question of differences in difficulty among rules, and the possibility that these differences may be attenuated through training.

Rule Learning

In Experiment I the observed improvement across a series of RL problems demonstrates that through training the subject does acquire knowledge about how a rule assigns stimulus patterns to response categories. At least for the rules which are initially least difficult, conjunction, inclusive disjunction, and joint denial, it has been shown that RL proceeds regularly to a level of mastery such that, given the relevant attributes, the subjects categorize stimulus patterns essentially without error.

Training on a single rule bears obvious similarity to the regimen of learning-set studies (Harlow, 1949). Solving each of a series of individual problems yields continuous improvement, producing optimal performance on a given type of task. In the present situation, where some rules are familiar and relatively easy at the outset, while others—because of lack of training or interfering expectations—are relatively difficult, this procedure operates to reduce or eliminate initial differences in difficulty. It may also be noted that the process called RL is not unlike concept formation, as that phrase is used by Piaget (1929; see also Flavell, 1963). When one of Piaget's children learns the concept of conservation of quantity, he has acquired, in effect, a general rule or way of responding in

a variety of problematic situations. For example, the attainment of conservation is implied when a child indicates that (a) emptying a container of beads into another of different shape and capacity does not affect the quantity of beads, or (b) different spatial arrangements of poker chips on a table do not entail a change in number of chips. In the last analysis, it is not the specific attributes, inextricably tied to the rule, which the subject acquires as the concept, but rather the rule itself.

Processes in Rule Learning and Identification

Rule learning may be construed as the process by which the subject acquires information on the assignment of all combinations of values on relevant stimulus dimensions to response categories. In the stimulus population used in the present experiments, this means a mapping of nine different patterns (combinations of three values on each of two relevant dimensions) onto a two-response system. Evidence from Experiment II, however, argues that RL introduces an additional, simplifying factor or heuristic into the categorization process. Simplification results from collapsing the 3×3 , in general $m \times n$, matrix (representing combinations of values on relevant dimensions) to a 2×2 matrix, each cell of which contains a set of patterns, called here contingencies, characterized by the presence or absence of the two relevant attributes. To say, then, that the subject has learned a rule is to say that he understands how it uniquely assigns contingencies to response categories.

It is conceivable that a subject who achieves, with practice, the capacity to react to all instances of each contingency as members of the same class might derive an algorithm for learning a new rule or for attaining solution to

any RI problem: *Observe the assignment of one example of each of the four contingencies, then classify all future instances according to this scheme.* This plan reduces a potentially unlimited population of instances to four, then merely maps these four onto a twofold response system. While it is probably true that this algorithmic process is not carried out formally by the subjects, it is not unreasonable to assume that many subjects perform a similar routine on a more intuitive level.

General RL, for at least some subjects, may be conceived, then, as a two-stage process involving: (a) reductive coding of a stimulus population to the matrix of attribute contingencies and (b) acquisition of the assignment of contingencies to response categories which is unique to each of a variety of rules. Neither the data nor this analysis indicate that Step *a* is mastered before the subjects learn anything about a rule. Both imply only that sophisticated and efficient performance in problems which require the learning of a new rule or the identification of a familiar one will include an intervening step which reduces the effective complexity of the stimulus population. Evidence from Experiment II, however, makes it clear that some subjects used the collapsing or coding heuristic prior to complete learning of the more difficult conditional and biconditional solutions. The similarity of the coding principle to the formal truth table is obvious. If the foregoing interpretation is correct, the results imply the utilization, as a mediator, of some informal version of this device, which of course is well known to have considerable power in logic.

Rule Difficulty

If the subject learns a rule as a specific and unique mapping of contingencies, and uses differences in the assign-

ment of contingencies as a means of identifying the rule (solution) for any RI problem, the implication is that all rules which in some way entail the combination of two (or any fixed number) relevant attributes are equally complex or difficult. It is well established that rules for combining two relevant attributes are not equally difficult at the outset even when the attributes are known and probably take different amounts of practice to learn. However, there is evidence in the present experiments that such differences may be reduced through RL and probably eliminated. Initial differences may be attributable to differing amounts of preexperimental experience with the various rules used here, and to peculiar, unfamiliar, or "unexpected" assignments of contingencies to response categories, such as the assignment of both TT and FF instances, which have no common relevant attributes, to the same response class. It is possible, moreover, that variation in the uncertainty of pattern-to-response or contingency-to-response assignments, e.g., Level III versus Level II rules (Neisser & Weene, 1962), may account for differences in initial difficulty. All of these sources of differential rule difficulty in RL apparently may be rendered ineffectual through training, leaving the subject with a repertoire of habits or strategies for solving RI problems with equal facility regardless of the rule. This repertoire may, in fact, be analogous to the knowledge of stimulus attributes which, it is assumed, the subject brings with him to most experimental situations.

Rule and Attribute Identification

Rule-identification tasks are analogous to those requiring identification of relevant attributes. In both, one aspect of the problem is unknown, either the rule (RI) or the relevant

attributes (AI). In both, given a complete understanding of the possible values of the unknown and a systematic plan or strategy for isolating the correct value(s), the subject may solve the problem in a computable minimal number of stimulus instances which depends almost exclusively on the number of possible values of the unknown.

The dichotomy implied by the terms RL and RI is debatable. The problem is that acquisition is in this case essentially a process of gradual change in performance with practice. It takes some arbitrary criterion to distinguish the attainment of habits from their utilization. One cannot appeal to superior performance during any so-called identification problem as evidence of a difference between learning and utilization of conceptual rules. Such superiority, which indeed exists, can be construed as reflecting in the identification phase merely a continuation of the improvement shown in the RL problems. The differentiation between RL and RI rests not on performance but in the nature of the experimental task. As the subject learns new rules, in fact, as soon as he has learned two, it is possible to present him with a problem requiring the identification of which of the known rules is correct. If the subject has not completely mastered either or both rules, some RL may occur. The degree to which the subject actually has mastered the set of rules is the degree to which any problem involves true RI.

It is difficult to specify the degree to which improvement in performance over a series of AI or RI problems results from increasing familiarity with the rule(s), or from the development and utilization of strategies appropriate to the identification of unknowns. Rules differ in familiarity to the unsophisticated subject (Wells, 1963), and it is virtually impossible to equate

them on this factor through preliminary instructions, no matter how detailed. Thus, "true" AI or RI conditions would necessitate thorough understanding of the relevant rule(s), such as might be provided by the RL conditions of the present experiments, prior to the presentation of identification problems. The effects of such prior training should be explored more thoroughly in subsequent experiments.

Theoretical Comment

Emphasis in this paper has been placed primarily on an analysis of conceptual problems; the result is as much a model of the task as it is a model of the organism. Even if the assumption is true that some subjects operate on a problem with an intuitive truth table, it holds only after considerable experience solving problems based on a variety of rules. Moreover, this assumption in no way generates a description of the manner in which initial rules are acquired, nor does it specify how unknown attributes are identified. An adequate theory of conceptual behavior must embody a set of principles which provides an empirically valid account of these processes.

The model of efficient performance in RL and RI tasks presented here presumes that the subject achieves an encoding of known *stimulus attributes* into an effective truth table; the unique distribution of the attribute contingencies of the truth table into categories then identifies the relevant rule. Most theoretical interpretations of conceptual behavior (e.g., Bourne & Restle, 1959; Restle, 1962) have been concerned primarily with a different case (AI) wherein it can be assumed that the subject knows the *rule* and must identify the relevant attributes only. Analytic separation of attribute- and rule-learning aspects of conceptual behavior does not preclude the development of

a theory capable of describing performance when both rule and attributes are unknown (CL). Yet, available evidence is clearly too fragmentary to sustain any set of strong theoretical assertions.

It may be noted that the RL model implies a hypothesis-testing strategy which yields a simple test among alternative possible solutions. Each new contingency reduces the number of possibilities until only the correct one remains. Restle (1962) and others have conceived of AI as proceeding along similar lines. That is, the subject selects, tests, and rejects stimulus attributes (as hypotheses) until the correct one or combination is discovered. In view of the considerable success of these models in providing detailed, quantitative accounts of performance in simple concept-identification tasks, it is reasonable to suggest an extension to more complicated CL problems.

If both the rule and the relevant attributes are unknown, the subject may (a) randomly select a pair of attributes on a provisional basis for encoding in a truth table, (b) test for systematic assignment of the attribute contingencies to response categories, and (c) reject the pair and select a new one if tests for consistent assignments fail. This extension implies a formal model with two independent processes, corresponding to discovery of the two unknowns. The processes obviously overlap, however, in the sense that the subject must test for systematic contingency assignments to response categories subsequent to the selection of any attribute pair.

It is clear that any such theory has little or no validity as an account of the behavior of naive subjects in CL tasks. Present data show that the subjects begin to encode relevant attributes only after an appreciable amount of practice on experimental problems,

and even then only when the relevant attributes are given at the outset. Thus, the accuracy of this description is very likely limited to those cases wherein the subject has attained *mastery* of several rules. Observations of the naive subject in the present experiment are closer to expectations based on Hunt's (1962) interpretation. Most subjects presume the problem to be conjunctive at the outset, with the obvious and predictable result that CL and AI conjunctive problems do not differ in difficulty. When the conjunctive strategy is insufficient, the subject discovers a set of partial solutions which, taken together, permit him to attain criterion. Thus, Hunt's model may provide an adequate description of the subject's performance on initial CL problems. While solution to such problems may be attained through the acquisition of a decision tree, as Hunt suggests, present data indicate that certain generic decisions in the tree—corresponding to the relevant rule—are transferable across problems, for evidence of learning to learn (or solve) is provided by both experiments. Since both inter- and intrarule transfer were observed, it seems likely that generic decisions somehow contribute to the formation of an intuitive truth table. The nature of this acquisition process is, however, obscure.

Suggestions and Conclusions

The analyses and results presented here immediately suggest certain problem areas requiring further study. First, and perhaps most obvious, is the exploration of parameters affecting RL and RI. Many of the variables which exert a pronounced effect in AI, such as number of irrelevant dimensions, may be expected to have little or no effect in RL. On the other hand, possibilities which are irrelevant to AI, such as the number of known rules,

may have direct and measurable effects in RI.

Second, there is a need for more rigorous study of AI under rules other than the conjunction. Consideration of previous results and the present analysis suggests that a large part of the differences across rules in previous experiments may be attributed to differential prior training and familiarization. Further study in which subjects are thoroughly trained on the relevant rule, so that it constitutes a true "given" in the problem, should clarify the question of inherent rule difficulty in AI. It may be anticipated that somewhat different strategies will be utilized when different rules are in effect, since it is apparent that those appropriate to conjunctive problems are not efficient for other rules (Bruner et al., 1956). Whether important determiners of conjunctive concept identification will have a different effect when other rules are used, however, is not clear. The results of Kepros and Bourne (1963) with biconditional problems suggest no difference in at least one case, that of increasing the number of irrelevant dimensions. But additional experiments to determine, for example, whether increasing the amount of intradimensional variability has a similar effect on AI with a variety of rules are needed.

Third, although previous work has shown that the proportion of positive instances has a significant effect on solving conjunctive problems (Hovland & Weiss, 1953), it is not safe to generalize this result to all concept types and problems, for it may be no more than an artifact of the appropriateness of a positive focus for conjunctive strategies (Hunt, 1962). Rules which are more easily solved with a negative focus may well show a significant superiority for negative instances. Apart from the question of strategies, the

present analysis suggests that the proportion of the four stimulus contingencies may be an even more powerful determiner of solution in certain problems than the raw proportion of positive and negative instances.

Fourth, the present experiments dealt only with rules based on the sentential calculus, i.e., rules reducible directly to expressions involving only negation, conjunction, and inclusive disjunction. Other classes of rules, expressing such relationships as equality, order, inclusion, and so forth, remain to be explored. It is of interest to observe how the subject determines the relevant relationship between given or known attributes, and to compare such performance with that on the logical rules used in the present study.

These and other problems remain to be investigated. Studies reported here, though only a beginning, attest to the advantages of separating important structural and behavioral components in conceptual problems. They indicate the obvious perhaps—that performance in these tasks is a complex process of thought and behavior, involving subtle mediating mechanisms which intercede between problem and solution. In addition, they provide some promising techniques for elucidating processes which may have been confounded by task requirements in previous research. There is much more to be accomplished, however, before a rigorous and complete understanding of these components, and the variables which govern them, is at hand.

REFERENCES

- ARCHER, E. J., BOURNE, L. E., JR., & BROWN, F. G. Concept identification as a function of irrelevant information and instructions. *Journal of Experimental Psychology*, 1955, 49, 153-164.
- BATTIG, W. F., & BOURNE, L. E., JR. Concept identification as a function of intra- and interdimensional variation. *Journal of*

- Experimental Psychology*, 1961, **61**, 329-333.
- BOURNE, L. E., JR. Effect of delay of information feedback and task complexity on the identification of concepts. *Journal of Experimental Psychology*, 1957, **54**, 201-207.
- BOURNE, L. E., JR., & HAYGOOD, R. C. The role of stimulus redundancy in concept identification. *Journal of Experimental Psychology*, 1959, **58**, 232-238.
- BOURNE, L. E., JR., & RESTLE, F. Mathematical theory of concept identification. *Psychological Review*, 1959, **66**, 278-296.
- BRUNER, J. S., GOODNOW, JACQUELINE J., & AUSTIN, G. A. *A study of thinking*. New York: Wiley, 1956.
- BULGARELLA, ROSARIA G., & ARCHER, E. J. Concept identification of auditory stimuli as a function of amount of relevant and irrelevant information. *Journal of Experimental Psychology*, 1962, **63**, 254-257.
- CONANT, M. B., & TRABASSO, T. Conjunctive and disjunctive concept formation under equal-information conditions. *Journal of Experimental Psychology*, 1964, **67**, 250-255.
- FLAVELL, J. H. *The developmental psychology of Jean Piaget*. New York: Van Nostrand, 1963.
- GARNER, W. R. *Uncertainty and structure as psychological concepts*. New York: Wiley, 1962.
- HARLOW, H. F. The formation of learning sets. *Psychological Review*, 1949, **56**, 51-56.
- HAYGOOD, R. C., & BOURNE, L. E., JR. Forms of relevant stimulus redundancy in concept identification. *Journal of Experimental Psychology*, 1964, **67**, 392-397.
- HOVLAND, C. I. A "communication analysis" of concept learning. *Psychological Review*, 1952, **59**, 461-472.
- HOVLAND, C. I., & WEISS, W. Transmission of information concerning concepts through positive and negative instances. *Journal of Experimental Psychology*, 1953, **45**, 175-182.
- HUNT, E. B. *Concept learning: An information processing problem*. New York: Wiley, 1962.
- HUNT, E. B., & HOVLAND, C. I. Order of consideration of different types of concepts. *Journal of Experimental Psychology*, 1960, **59**, 220-225.
- HUNT, E. B., & KREUTER, JANET M. *The development of decision trees in concept learning: III. Learning the connectives*. Los Angeles: Western Management Sciences Institute, 1962.
- KEPROS, P. G., & BOURNE, L. E., JR. The identification of biconditional concepts. *American Psychologist*, 1963, **17**, 424. (Abstract)
- NEISSER, U., & WEENE, P. Hierarchies in concept attainment. *Journal of Experimental Psychology*, 1962, **64**, 640-645.
- PIAGET, J. *The child's conception of the world*. New York: Harcourt, Brace, 1929.
- QUINE, W. V. O. *Mathematical logic*. Cambridge: Harvard Univer. Press, 1958.
- RESTLE, F. The selection of strategies in cue learning. *Psychological Review*, 1962, **69**, 329-343.
- SHEPARD, R. N., HOVLAND, C. I., & JENKINS, H. N. Learning and memorization of classification. *Psychological Monographs*, 1961, **75** (13, Whole No. 517).
- TRABASSO, T., & BOWER, G. Concept learning in the four-category concept problem. *Journal of Mathematical Psychology*, 1964, **1**, 143-169.
- WALKER, C. M., & BOURNE, L. E., JR. Concept identification as a function of amounts of relevant and irrelevant information. *American Journal of Psychology*, 1961, **74**, 410-417.
- WELLS, H. Effects of transfer and problem structure in disjunctive concept formation. *Journal of Experimental Psychology*, 1963, **65**, 63-69.

(Received May 18, 1964)

METHOD FOR CHANGING STEREOTYPED RESPONSE PATTERNS BY THE INHIBITION OF CERTAIN POSTURAL SETS¹

FRANK PIERCE JONES

Institute for Psychological Research, Tufts University

An empirical method is described for changing habitual response patterns by inhibiting postural sets which disturb the reflex balance of the head. The procedure results in a redistribution of postural tonus which is reported by S as a decrease in the feeling of weight and in the effort needed to move. Differences in posture and movement are recorded by multiple-image photography, X-ray photography, and electromyography. Anatomical and physiological mechanisms are suggested to explain the phenomenon. The implications for behavioral science are discussed.

The problem of behavioral change becomes more important as society becomes more complex. Even if a complete set of desirable responses could be successfully taught to everyone, the necessity would still remain, in a rapidly changing world, of modifying or eliminating some of them. This is easier said than done, however, as anyone can testify who has attempted to change an unwanted habit. A well-learned response pattern continues to

be elicited by appropriate stimuli long after it has lost all value for the organism. "I see the better course and approve of it," the poet says, "but I follow the worse."

This paper will describe an empirical method for changing habitual patterns of behavior by inhibiting certain postural "sets." Once understood, the method can be applied to any pattern of response. It is demonstrated most easily with a movement which involves head, neck, and trunk and which is performed against gravity.

POSTURAL SET

The term *set* is used here to mean a preliminary change in the level and distribution of tension as a preparation for movement (Jones, 1963). In preparation for a movement against gravity, the increase in tension is often sufficient to change the relation between head and trunk. This change is so much a part of the movement to come that it is seldom detected by the mover. If it is eliminated from the response pattern (i.e., inhibited), the movement itself will differ markedly from the ordinary movement in the

¹ The research reported in this paper was supported by grants from the Carnegie Corporation of New York and the United States Public Health Service (GM-04836), with the help of the Tufts Fund for Research in Kinesthesia and the Katherine Bowditch Codman Fund.

This study originated in conversations with J. Dewey, who persuaded me that the sensory phenomenon described in this paper could (and should) be investigated within the framework of experimental psychology. The study was undertaken with the advice and encouragement of G. P. McCouch, J. McV. Hunt, and H. Schlosberg.

For assistance in the design and conduct of experiments and the interpretation of experimental data I wish to thank my colleagues at the Institute for Psychological Research, particularly M. and Dorothea Crook, Florence E. Gray, and J. A. Hanson.

way the weight of the body is perceived. In addition, the movement will produce a different distribution of postural tonus, against which a subsequent set can be perceived kinesthetically as a figure-ground relation. The character and significance of the phenomenon can perhaps be best understood from an introspective account.

INTROSPECTIVE REPORT

The effect on a movement pattern of changing the habitual relation between head and trunk was first demonstrated to me by the late A. R. Alexander.² Alexander used the movement from sitting to standing for the demonstration. The experience, which is still vivid in retrospect, was unexpected. Assuming that the demonstration would have something to do with voice production, I did not anticipate a movement of any kind. At the start, Alexander made a few slight changes in the way I was sitting. These changes seemed quite arbitrary to me, and I could not remember afterwards what they were. Then, asking me to leave my head as it was, he initiated the upward movement without further instructions. The movement (from sitting to standing) was completed before I had a chance to organize any voluntary response. The sensation was not that of either getting up or being lifted. My body seemed to be

straightening reflexly against gravity. The analogy of the knee jerk suggested itself immediately. The two movements were similar kinesthetically, but this one, in an integrated fashion, involved the whole extensor system.

In addition to the reflex effect, the movement was notable for the way time and space were perceived. Though less time than usual was taken to complete the movement, the rate at which the head and other parts moved was paradoxically slower, and the trajectories which they followed were unfamiliar. The experience is difficult to describe, but the impression was that of a sudden expansion in both dimensions so that more time and space seemed available for the movement.

The sense of reduced weight which accompanied the movement persisted after it was completed. The centers of gravity of the head and trunk seemed to have shifted forward in relation to the feet, and the weight of the body to be supported by structures which previously had not been called into play. The effect of the change was to eliminate much of the fatigue which for me had been associated with standing.

What later became a clear-cut, easily recognizable kinesthetic experience of lightness and ease of movement registered at first as the absence of sensations which, though familiar, had never been consciously observed. The nature of these sensations began to be manifest, however, when an attempt was made to repeat the movement from sitting to standing. This time I got set in anticipation. There was an increase of tension in the neck, trunk, and limbs as soon as the suggestion of getting up was made. The tension had apparently always been present as a preliminary to movement, but by some process of adaptation it had passed unobserved. Now, against the remembered back-

² A. R. Alexander was a brother of F. M. Alexander (1923, 1932, 1941). Alexander's teaching, which was based on a principle of conscious inhibition, had an important influence on the thinking of Dewey (1922, 1923, 1932; McCormack, 1959). Coghill (1941) and Dart (1947, 1950) gave firsthand accounts of Alexander's system. It was noted briefly but favorably by Sherrington (1946) and Herrick (1949) and discussed at various times in the British medical press (Barlow, 1945, 1954; Macdonald, 1926; Rugg-Gunn, 1940). A discerning account has recently been written by Carrington (1963).

ground of the previous movement, it stood out as a recognizable pattern. It was as if the reflex movement had eliminated some of the "noise" from the system, so that when the signal reappeared it could be perceived as a discrete pattern of spreading tension. The pattern began first in the neck, fixing the position of the head and spreading quickly to the trunk and limbs. While it was present, no movement except the habitual seemed possible.

Once this signal—this pattern of spreading tension—had been clearly perceived, I found it reappearing in a great variety of situations—in speaking, for example, in taking a deep breath, in climbing stairs, in lifting a heavy weight, in changing the fixation of the eyes. The pattern was magnified by stress, but seemed to be present in some degree a great deal of the time. It was repeatedly demonstrated to me that when this pattern was inhibited and the tension prevented from spreading, the character of the associated activity changed: it became markedly easier and was accompanied by the same kinesthetic effect of lightness I had observed before.

The kinesthetic effect of lightness was tonic and persisted long after a particular movement had been completed. With it almost from the start came certain automatic or semiautomatic changes which are probably significant for understanding the phenomenon: (a) a change in the rest position of the eyes and a marked reduction in the tension used in eye movements; (b) a change in the rest position of the jaw and a relaxation of tension in the tongue and throat; (c) a change in the rate and depth of breathing, associated with an increase in the excursion of the diaphragm. All of these changes were corollaries of the postural change and disappeared

when I returned to my habitual pattern of posture and movement.

I have described a clear-cut sensory experience. Though kinesthetic in origin, it had the sharp and immediate reality of experience obtained through other sense modalities. It was repeatable, and it was not accompanied by any detectable form of suggestion. I was immediately struck by the implications of the experience for behavioral science. It seemed to open up to controlled investigation an area of the self which has heretofore been notoriously private and inaccessible and to provide the individual with a reliable means for changing a behavior pattern without delving into the past or having recourse to some outside authority. The inhibitory control which I found myself using in simple, everyday movements seemed capable of extending to any situation which could be expressed in terms of stimulus and response. Before speculating, however, about implications and extensions, I decided that experimental data were needed in order to define the phenomenon operationally. To this end, a series of studies have been made at the Tufts Institute for Psychological Research. I shall review them here and advance a theory of mechanism to account for the data.

PROCEDURES FOR CHANGING THE BALANCE OF THE HEAD³

The subject, seated in a comfortable, "relaxed" posture, is asked to straighten up into a more erect posture. As he responds, a slight change can be seen to take place in the axis of his head, which is usually rotated backward, bringing the occiput closer to the seventh cervical vertebra. If he is asked to straighten up even further

³ The procedures described here are adapted from procedures used by F. M. Alexander.

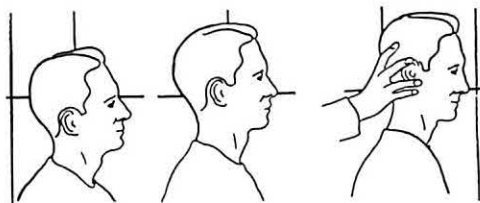


FIG. 1. Three sitting postures, traced from photographs. (From left to right: habitual relaxed, habitual erect, and experimental. In the final picture the subject has been guided into the erect posture by the experimenter, who counterbalanced the backward pull on the subject's head during the movement.)

into his "greatest sitting height," the angle of head rotation will increase (Jones, Gray, Hanson, & Shoop, 1961). After the subject has returned to his relaxed posture, the experimenter applies a light pressure at the back of his head just sufficient to counterbalance the backward rotation. The subject is then guided in a series of small movements, while the experimenter continues to prevent the backward rotation of his head, taking care not to evoke stretch reflexes in neck muscles by applying unnecessary pressure. At the end of the procedure, the subject will again be in a more erect posture, but the relation of head to trunk will be different from that observed in the habitual posture (see Figure 1).

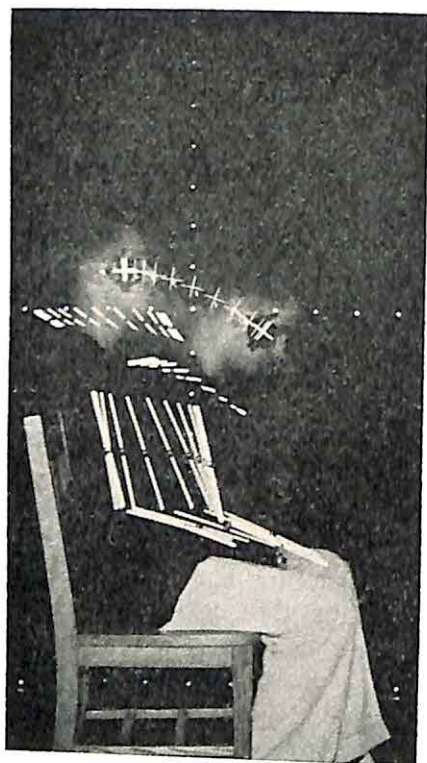
If these procedures are carefully followed, the subject's weight will seem to the experimenter to be progressively reduced until little effort is needed to move him passively. Any movement carried out after the change will follow a different course from the corresponding habitual movement.

EFFECT OF HEAD BALANCE ON MOVEMENT PATTERN

The effect of a change in head balance on movement pattern was first demonstrated by Jones and Narva (1955) with multiple-image photog-

raphy. Small electric lights were attached to the subject's head, trunk, and limbs and connected with wires to a battery. The subject moved with profile to the camera, which was left open throughout the movement while an episcotister (or "Marey wheel") interrupted the moving image 10 times a second. The images of the interrupted lights provided a time-space pattern of the movement. Though the technique lacked precision, the photographs showed convincingly that the pattern of a movement is altered when the head balance is changed.

Jones and O'Connell (1956) refined the method by attaching strips of Scotchlite reflecting tape to the subject and recording the moving image by repetitive strobe at rates of 5, 10, and 20 flashes per second. The stroboscopic method is highly flexible and allows the subject a maximum of freedom to move. Patterns obtained in this way are sharp and clear-cut and contain a vast amount of information that can be quantified. An example of the method is shown in Figure 2. The subject wears a black jacket to cut down reflectance. A small cross, which is centered in the Frankfort plane (Howells, 1937) halfway between the tragon of the ear and the lowest point of the orbit, marks the position of the subject's head and its angle of rotation. Other markers are attached over the seventh cervical vertebra and the sternal notch and to the upper and lower arms. In the picture on the left, the subject sits leaning forward in the chair and moves back into an upright sitting posture. The strobe is pulsed at 5 flashes per second. In the picture on the right, the movement is repeated after the experimenter has changed the relation between the subject's head and trunk. In the experimental movement, the trajectory of the head is higher, the



2a. Habitual.



2b. Guided.

FIG. 2. Stroboscopic multiple-image photographs, from leaning forward to sitting erect. (Reflecting markers are placed on: head; neck, at seventh cervical vertebra; chest, at sternal notch; upper and lower arm. Strobe at 5 flashes per second.)

curve is smoother and more regular, and the change in head pattern is accompanied by changes in the patterns of the trunk and arm.

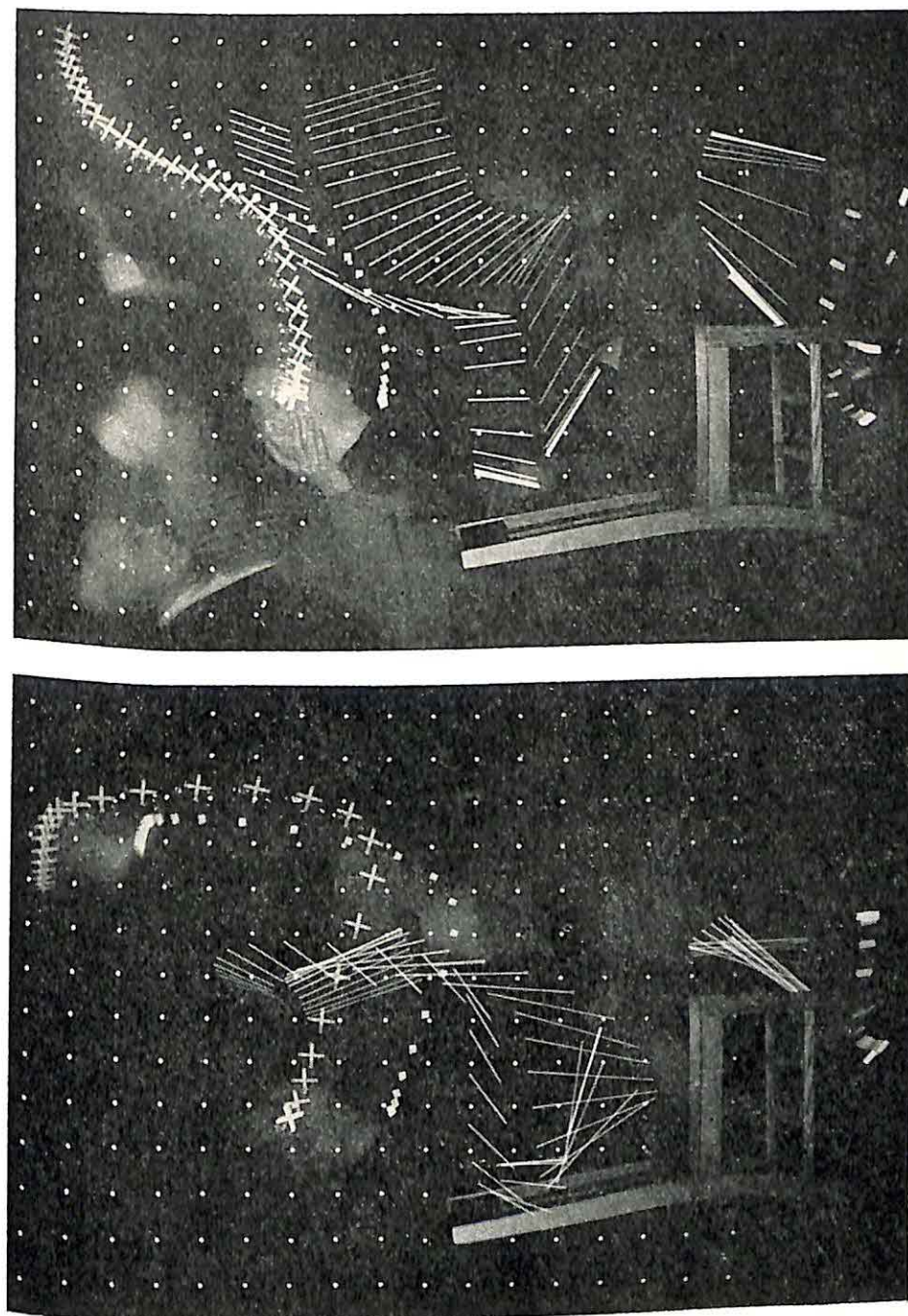
Other everyday movements which have been recorded in this way are walking, stair-climbing, stooping down to pick up an object from the floor, straightening up from a slump, and the movements from sitting to standing, from standing to sitting, and from lying down to sitting up. In each of these movements, the pattern is characteristically altered when the relation of head to trunk has been changed.

MOVEMENT FROM SITTING TO STANDING

The most striking differences in pattern are seen in the movement from

sitting to standing (see Figure 3). In the picture on the left, the subject, with markers on head, neck, upper and lower arm, leg, and foot moves from sitting to standing in his habitual way. In the picture on the right, the movement is guided by the experimenter, who has changed the relation between head and trunk in the manner described above.

In the movement from sitting to standing, habitual patterns show individual differences which remain remarkably constant from trial to trial (Jones & Hanson, 1961, 1962). When the movement is guided, individual differences tend to disappear so that the guided patterns are much more uniform than the habitual. Quantitative methods of pattern analysis (Jones &



3a. Habitual.

3b. Guided.

FIG. 3. Stroboscopic multiple-image photographs, from sitting to standing. (Reflecting markers are placed on head, sternal notch, right upper and lower arm, right leg, right foot, and left foot. Strobe at 10 flashes per second.)

O'Connell, 1958) were used to compare guided with habitual movements. In an analysis of 36 patterns from six subjects, indexes were found that distinguished the guided from the habitual at a confidence level of .01 or better (Jones, Gray, Hanson, & O'Connell, 1959). The three indexes which have proved most useful in subsequent studies are: (a) head thrust, which measures the forward thrust of the head during movement; (b) trajectory ratio, which measures the extent to which the head trajectory departs from a straight line; and (c) rise time, which measures the time needed to bring the head above the starting level. In the guided movement, head thrust decreased, rise time was shorter, and head trajectory approached more closely a straight line.

The value of the three indexes was tested by applying them to patterns which could be judged "better" or "poorer" on the basis of some external criterion. Jones and Hanson (1961) analyzed the movement from sitting to standing as it was performed by "well-coordinated" and "poorly coordinated" subjects. Jones, Hanson, Miller, and Bossom (1963) compared the same movement in a group of normals and a group of patients with neurological diseases. The three indexes which had distinguished guided movements from habitual distinguished (again at a confidence level of .01 or better) well coordinated from poorly coordinated and normal from abnormal movements. By these criteria, then, the guided movements are not only different from the habitual, they are better.

SUBJECTIVE EXPERIENCE

For the great majority of subjects, the guided movements are accompanied by a kinesthetic experience of lightness and ease of movement, which tends to persist after the experimental session

is concluded (Jones, 1954). By some, this kinesthetic effect is observed immediately, i.e., after the first guided movement. Others observe it only after the movements have been repeated a number of times. When asked to describe the experience, subjects are apt to find it difficult to put their feelings into words. To make the task easier, a check list of 18 adjectives was constructed, 16 of them paired opposites (Jones, 1964). The subject is first asked if the guided movements felt different in any respect from his ordinary movements. If the answer is "yes," he is given the list of adjectives and asked to check those that best describe the difference.

Data from 39 subjects are presented in Table 1. All of the subjects were normal young adults without previous knowledge of the phenomenon under study. The responses were obtained after a brief demonstration of 10 or 15 minutes.

The adjectives most frequently checked (i.e., lighter, less familiar, higher, and smoother) seem to give the core of the sensory experience. Other adjectives reflect individual differences in the attempt to describe the unfamiliar. A feeling of "unreality" which some subjects observed was variously identified by "brighter," "duller," or "softer." The fact that some subjects checked "slower," while others checked "faster," seems related to the paradoxical character of some of the guided movements (e.g., sitting to standing) which may be completed in a shorter time than the corresponding habitual movements, but which reach a lower maximum velocity. Sometimes the experimenter failed to convey the kinesthetic effect of lightness. When this was the case, a different constellation of adjectives was checked with "heavier" and "more difficult" replacing "lighter" and "easier." Some sub-

TABLE 1
PERCENTAGE OF RESPONSES OF 39 SUBJECTS TO THE ADJECTIVE CHECK LIST

Adjective	%	Adjective	%
Lighter	72	Tenser	20
Less familiar	62	Brighter	15
Higher	59	More difficult	15
Smoother	54	Less steady	13
Slower	44	Heavier	13
More relaxed	44	Faster	10
Easier	41	Jerkier	10
Softer	38	Duller	8
Steadier	36	Lower	3

jects reported that the feeling of lightness was confined to the upper part of the body at the start and did not extend below the waist until a later stage in the demonstration, or until a later session. The spread of the kinesthetic effect from the head downwards (it is never reported the other way around) is undoubtedly important for understanding the phenomenon.

To measure the subjective feeling of reduced weight, the method of magnitude estimation has been used. The subject is asked to put a value of 10 on the effort he ordinarily exerts in a movement against gravity; then, on the same basis, he is asked to estimate the effort involved in the experimental movements. Measured in this way, the feeling of weight may be reduced anywhere from 20 to 80%, with some subjects reporting that all sense of effort has disappeared.

X-RAY STUDY OF THE HEAD-NECK RELATION

X-ray photography was used by Jones and Gilley (1960) to obtain a better definition of the three postures—habitual relaxed, habitual erect, and experimental—shown in Figure 1. The subjects were 20 students at the Tufts University School of Dental Medicine who were taking part in a study of dental occlusion. X-ray pho-

tographs were taken of each of them in the three postures. To analyze the photographs, planes were constructed through the head and neck, and the angles which these planes made with the horizontal and with each other were measured. Additional measures were taken of the distance between the spines of the first two vertebrae and of the distance between *sella turcica* (which corresponds roughly to the center of gravity of the head) and a horizontal line drawn through the second vertebra. When the three postures were compared, it was found that both of the erect postures differed from the habitual relaxed in the angle between horizon and neck. They differed from each other in the angle between horizon and head, and in the angle between head and neck. In the experimental posture, the distance between the first two vertebrae was greater, and the distance between *sella turcica* and the second vertebra was smaller, than in the habitual erect. These differences are significant at the .01 level or better.

X-ray photographs of the two erect postures are shown in Figure 4. In the experimental posture (on the right) the neck has lengthened, the head has rotated forward on the atlas, the distances between the spines of the vertebrae and between the vertebrae

and the occiput have increased. Though the head is slightly higher in the experimental posture than in the habitual, its center of gravity relative to the vertebrae is lower.

ROLE OF THE STERNOMASTOID MUSCLE

In the light of the X-ray findings, it seemed reasonable to expect a difference in the activity of neck muscles in the two erect postures. Preliminary investigation pointed to the sternomastoids (the prominent diagonal muscles on either side of the neck) as the muscles which distinguished most sharply between the two postures. Later, the postural activity of these muscles was systematically studied in seven male subjects between the ages of 16 and 21 (Jones, Gray, Hanson, & Shoop, 1961; Jones, Hanson, & Gray, 1961). Surface electrodes were placed over the right sternal head halfway between origin and insertion. The ac-

tivity of the muscle was recorded as the subject sat in his "most comfortable" posture, in his "best" posture, and at his "greatest height." It was recorded again when he was guided into the "experimental" posture. Mean potential for the seven subjects increased from 7 microvolts for most comfortable to 10 microvolts for best and to 26 microvolts for greatest sitting height. In the experimental posture it dropped back to 6 microvolts. The difference in the behavior of the sternomastoid appeared not only in the postures themselves but also in the movements which led into them. All differences (which are illustrated by bar graphs in Figure 5) are significant at a confidence level of .05 or better.

ANATOMICAL MECHANISMS

The head can be rotated, tilted, or lowered by contracting the muscles attached to it; but it cannot be lifted. It can be lifted only by straightening and

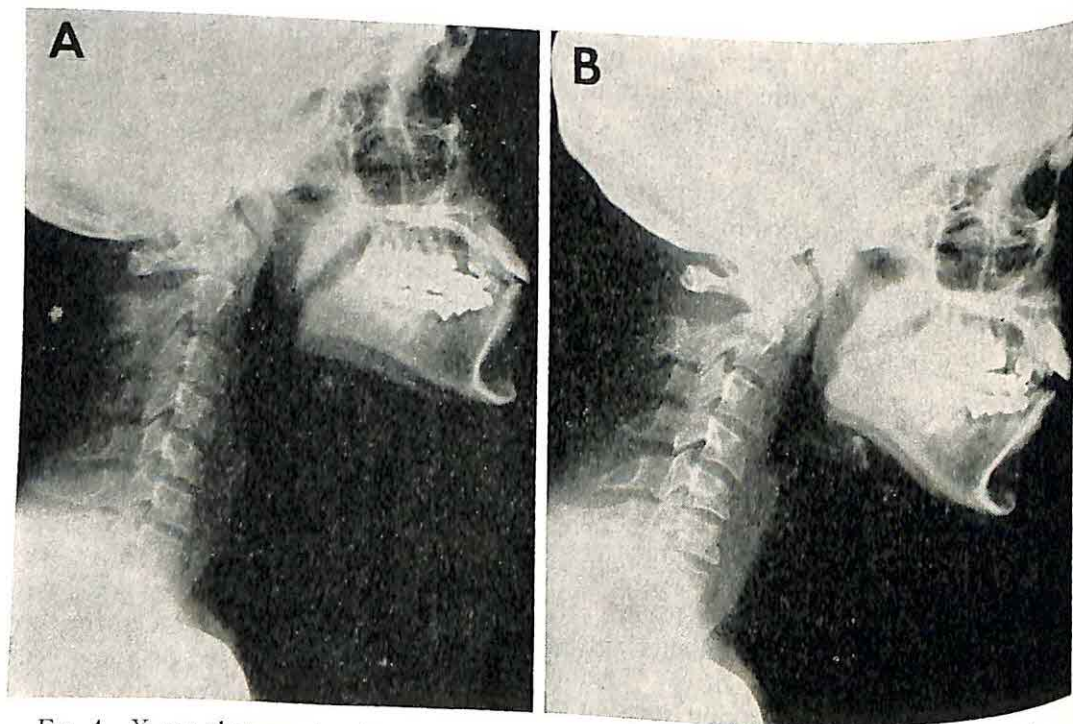


FIG. 4. X-ray photographs of two erect sitting postures. (A, habitual; B, experimental.)

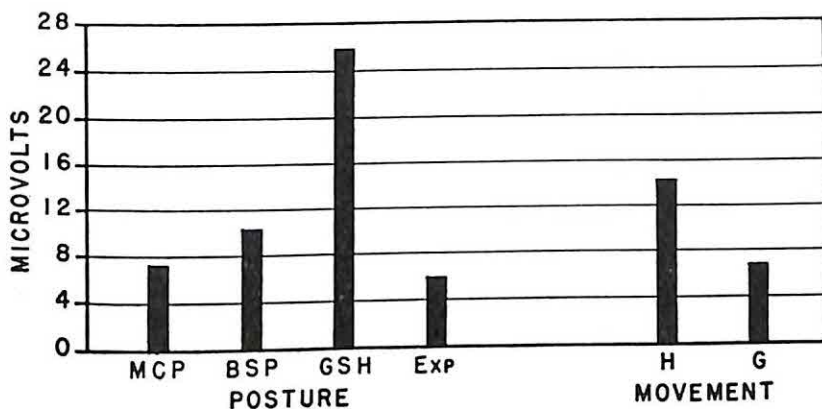


FIG. 5. Postural activity in right sternomastoid muscle. (Means for seven subjects. Abbreviations are: MCP, most comfortable sitting posture; BSP, best sitting posture; GSH, greatest sitting height; Exp, experimental; H, habitual movement into erect sitting posture; G, guided movement into erect sitting posture.)

lengthening the cervical spine. Between the vertebrae and the skin there are four structures whose joint action determines the height of the head and its angle of tilt and rotation. They are shown schematically in Figure 6.

Intervertebral Discs

The cartilaginous discs between the bodies of the vertebrae contain fluid and are capable of exerting hydraulic force on the bony structures around them (Gray, 1942, p. 281). The discs are kept under pressure by ligaments and by various muscles, including the small muscles which run from one vertebra to the next. If these muscles shorten, the distance between the vertebral bodies will be lessened, and the discs will be further compressed. Conversely, if they lengthen, the distance will increase as a result of the released pressure of the discs. In the intervertebral discs, then, is a mechanism by which the height of the head can be altered a small but significant amount.

Flexors and Extensors of the Neck

Both convexity (extension) of the cervical curve and its forward inclination (flexion) are countered by the

tension of ligaments and muscles whose origins and insertions are on the vertebrae themselves (Gray, 1942, pp. 390, 394-395). Acting together, the flexors and extensors of the neck straighten and strengthen the cervical spine, turning it into a column of support and, in so doing, increase the height of the head.⁴

⁴ In standard textbooks of anatomy, the extensors of the head and the extensors of the neck are treated under a single heading and illustrated with the same plates so that it is sometimes difficult to distinguish one group of muscles from the other. They have similar names (*splenius*, *longissimus*, etc.) and similar origins on the spinal column. They differ, however, in their insertions and in the functions which they perform. When the neck is flexed, the head is brought forward and down. If, then, the neck extensors contract, the spine will be straightened (extended), and in the process the head will be brought to a higher level, even though the head extensors remain relaxed. Contracting the head extensors, on the other hand, will not lift the head, but will only tilt it backward so that the occiput approaches the seventh cervical vertebra. Duchenne (1959, p. 534) first pointed out in 1867 the difference of function between the two sets of extensor muscles. He had observed that patients whose neck extensors were paralyzed could not lift their heads, though they still had the full use of their head extensors.

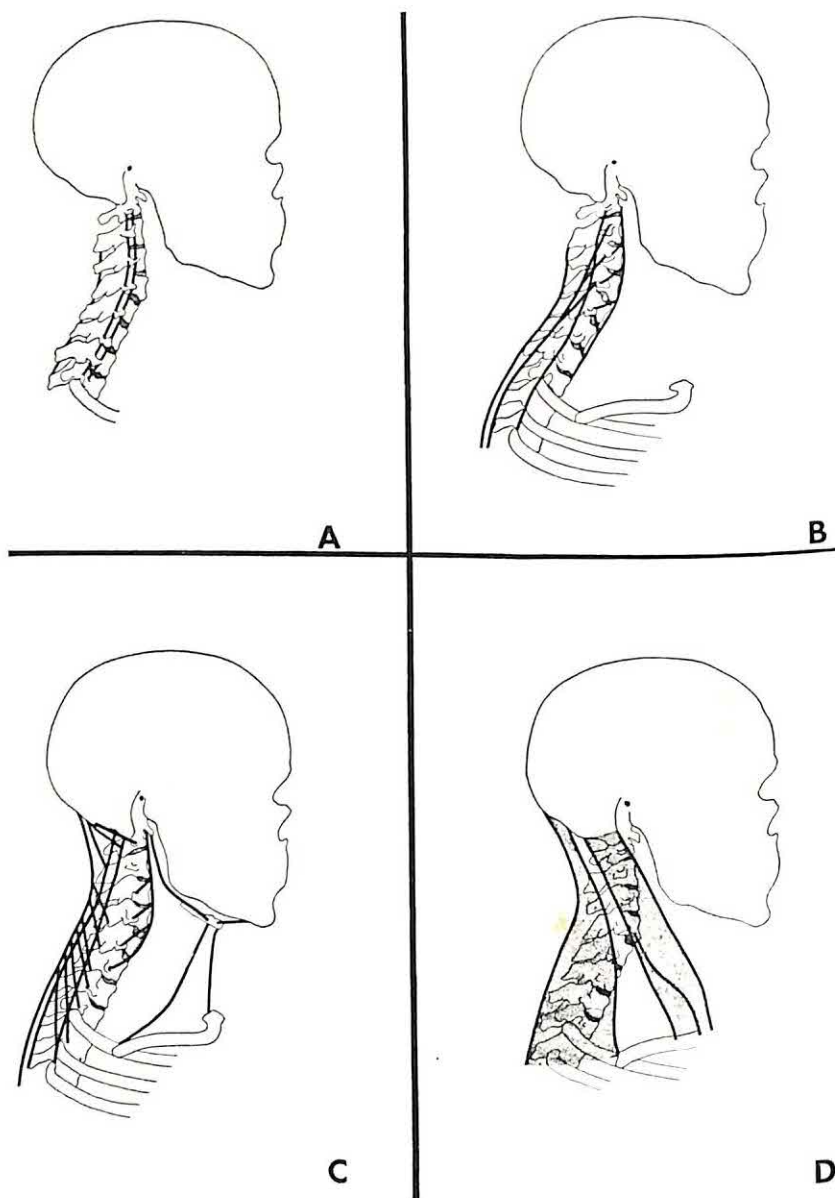


FIG. 6. Anatomical structures affecting the height and angle of the head. (A, intervertebral disks with *interspinales* and *intertransversarii* muscles. B, flexors and extensors of the neck. C, flexors and extensors of the head. D, *upper trapezius* and *sternocleidomastoid* muscles.)

Flexors and Extensors of the Head: I

In the upright posture, the head is in unstable equilibrium, its center of gravity forward of the base of support (the atlas), and its weight balanced by the tension of ligaments and muscles which run from insertion points at the

base of the skull to various origins along the spinal column. These muscles, the extensors of the head, are designed to balance, in addition to the weight of the head, the tension in a smaller group of flexor muscles in front. By the lengthening and short-

ening of the various muscles in these two groups the head can be moved in many directions without destroying the equilibrium of forces (Gray, 1942, pp. 386-395).

Flexors and Extensors of the Head: II. Sternomastoid and Upper Trapezius

There are two pairs of flexors and extensors which connect the head directly with the shoulder girdle. They are the *sternocleidomastoids* (Gray, 1942, p. 384) and the *upper trapezii*, which form a blunted isosceles triangle at the back of the neck (Gray, 1942, p. 428). Simultaneous contraction of the two pairs of muscles will bring the head closer to the shoulder girdle and increase the curve of the cervical spine, the sternomastoid drawing the head forward and down, the *upper trapezius* preventing forward rotation by retracting the occiput.⁵ In the process, the origins and insertions of the muscles and ligaments which maintain the weight of the head against gravity will be brought closer together.

STARTLE PATTERN

An example of the joint action of the sternomastoid and *upper trapezius* can be seen in the "startle pattern" of Figure 7. Surface electrodes have been attached to the subject's skin over the right sternomastoid and right *upper trapezius*. In Figure 7A he is standing in his "most comfortable posture." In Figure 7B he has been startled by the sudden slamming of a door. The two sets of muscles have contracted simultaneously. The head

is thrust forward, but the Frankfort plane remains horizontal. The postural change does not stop with the head and neck. As in one of the "attitudinal reflexes" described below, the shoulders are lifted, the chest is flattened, the legs are flexed, and the arms are extended.

The startle pattern, which was studied with high-speed photography by Landis and Hunt (1939), provides a vivid example of how "good" posture can change to "bad" in a very brief time. The pattern itself is typical of bad posture in general, whether it is the result of age, disease, or lack of exercise (Jones, Hanson, & Gray, 1964). In the startle pattern, the active character of malposture and the sequence of events by which it comes about can be clearly observed. The response is not instantaneous. It begins in the head and neck, passing down the trunk and legs to be completed in about $\frac{1}{2}$ second. The neck muscles are central in the organization of the response. Jones and Kennedy (1951), who studied the startle pattern by multiple-channel electromyography, placed surface electrodes in various locations on the subject's neck, trunk, and limbs, with one pair always over the *upper trapezius*. The intensity of the stimulus was varied from the sound of a dropped book to the sound of a .32 caliber revolver. Sixty patterns were obtained from eight subjects. In all cases when the stimulus was strong enough to elicit a response, it appeared in the neck muscles; in many cases, the response appeared nowhere else.

PHYSIOLOGICAL MECHANISMS

I have described two reciprocal sets of structures in the neck, one designed to straighten the cervical spine and move the head out from the trunk, the other designed to bring head and

⁵ It may be significant that the two muscles have the same phylogenetic origin (Romer, 1949, p. 285) and that, unlike other neck muscles, they receive their principal innervation from a cranial nerve (the accessory). Duchenne (1959, p. 5) commented on their extreme sensitivity to electrical stimulation, which he attributed to the peculiar character of their nerve supply.

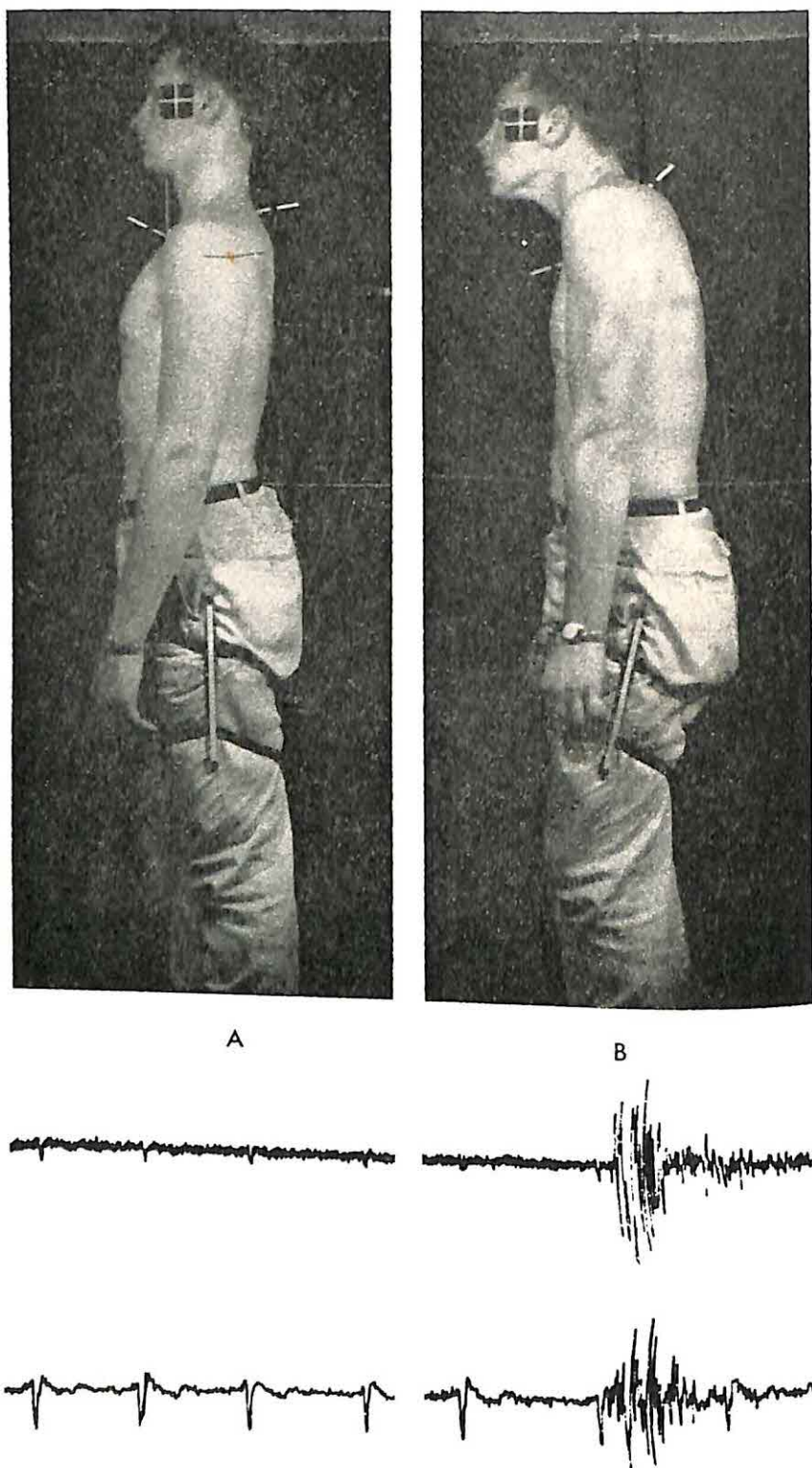


FIG. 7. Activity of neck muscles in the startle pattern. (A, subject standing in "most comfortable posture." B, subject after door slam. EMG records: above—*upper trapezius*; below—*sternomastoid*.)

trunk closer together and increase the spinal curvature. In the procedure described earlier in this paper, the downward pull of the sternomastoid and *upper trapezius* was counterbalanced by the experimenter while the subject moved. It is conjectured that in the process, the pressure on the discs was reduced, the flexors and extensors of the neck shortened, and a lengthening took place in the flexors and extensors of the head and in the small muscles between the vertebrae. During movement, the lengthening which appeared in the neck could be expected by purely mechanical means to extend along the rest of the spine, with the force being transmitted by muscle and ligament from one vertebra to the next.

Stretch Reflexes

The mechanical effect of stretch on muscle is modified by the action of the nervous system. When a muscle is stretched, it contracts reflexly, and the strength of the contraction is proportional to the stretch. Though stretch reflexes occur at the spinal level, they are under the influence of higher centers in the nervous system. This influence is transmitted directly to the muscles through the system of small motoneurons, the gamma efferents. As the level of activity in the gamma efferents rises or falls, the sensitivity of a muscle to stretch increases or decreases. Whether a contracting muscle shortens or lengthens (as in lowering a weight), the stretch reflex persists as long as the stretch stimulus is applied. Though the range of lengths over which muscles can contract is considerable, there is an optimal "resting length" at which a particular muscle exerts maximum tension.

It seems reasonable to suppose that one effect of the experimental pro-

cedures is to reorganize stretch reflexes in the neck and back so that more muscle fibers take part in the antigravity response, and that the muscles in contracting remain closer to their optimal resting length. This effect would account for the subjective feeling of reduced weight while maintaining a position as well as for the apparent lack of effort in movements performed against gravity.

Head-Neck Reflexes

The kinesthetic phenomenon and the change in the movement pattern cannot be explained entirely by the behavior of individual muscles. Such an explanation does not account for the major role which the head and neck appear to play in facilitating the movement, nor for the secondary changes, e.g., in breathing, reported by subjects.

In the posture of animals, it is well established that the most important mechanism of control is the head-neck relation. Magnus and his associates (Magnus, 1924; Magnus & de Kleijn, 1912; Rademaker, 1931), in a long series of carefully controlled experiments, showed that the position of an animal's head in space or its position relative to the body can affect the distribution of tonus in its neck, back, and limbs. Magnus summed up the principle by saying that in posture and movement the head leads and the body follows.

In classifying his material, Magnus used two categories—the "attitudinal reflexes" and the "righting reflexes."

The *attitudinal* reflexes with receptors in the labyrinths and in the joints of the neck (McCouch, Deering, & Ling, 1951) are used by an animal to maintain a position that is assumed for some special purpose, like that of a cat drinking from a saucer or looking up at a piece of meat held high. The position taken by the head imposes on

the rest of the body an attitude which is maintained as long as it is functional. These attitudes are quite stereotyped and follow regular rules. According to Magnus, they are the most enduring and untiring of reflexes.

The *righting* reflexes take over when an animal is ready to return to the normal upright posture. The fixed position of the head is released; the imbalance of parts, registering at a postural center in the brain, initiates the righting response; as in the attitudinal reflexes, the head leads and the body follows. The mechanism is most strikingly demonstrated when a cat is held on its back in the air and then dropped. The instant it is let go it begins to right itself. The head turns first. As it turns, the tensions in the neck, back, and limbs are progressively altered. The body is twisted around, and the cat lands on its feet. (Marey's photographs of the phenomenon are reproduced by Chatfield, 1957, p. 205.)

Secondary effects of the head-neck reflexes on respiration, circulation, and eye position have also been demonstrated. In diving birds and mammals there is a highly developed set of reflexes which stops breathing and slows down the heart in order to conserve oxygen (Irving, 1939; Scholander, 1963). In ducks, this reflex is strongly reinforced by the change in the relation of head, neck, and body that takes place in diving. Huxley (1913) found that without immersing a duck in water, breathing could be stopped by bringing the head and neck into the diving position or by placing the duck on its back and dorsiflexing the head. Conversely, breathing was at once restored if the head was brought back to the normal posture. A similar postural mechanism to stop breathing and slow the heart was later demonstrated in diving mammals (Koppányi, 1929).

A. de Kleijn (1920) used monkeys to demonstrate the effect of head-neck reflexes on the position of the eye in the orbit. He showed that the eyes moved down when the head was dorsiflexed and up when it was ventrified, and that the same shift in eye position took place when the head remained fixed and the trunk was rotated so as to bring the back closer to the occiput or the chest closer to the chin.

Head-Neck Reflexes in Man

In human beings, the influence of the head-neck reflexes is masked by patterns of voluntary activity. The mechanisms are clearly present, however. They have been frequently demonstrated in infants (Gesell, 1954; Peiper, 1963),⁶ young children (Landau, 1923; Schaltenbrand, 1925), patients with neurological diseases (Simons, 1923; Walshe, 1923), and in normal adults (Hellebrandt, Schade, & Cairns, 1962; Tokizane, 1951; Wells, 1944). A large number of drawings and photographs to illustrate the patterns of head-neck reflexes as they manifest themselves in dancing, sport, and everyday activity were brought together by Fukuda (1957).

Various specific reflexes, both attitudinal and righting, have been defined in the literature. Rather than describe them in detail, I should like here to emphasize what each set of reflexes has in common. In the attitudinal reflexes, the head is drawn into a fixed position and tonus is redistributed in the trunk and limbs. In the righting reflexes, again under the influence of the head, the normal distribution of tonus is restored. These two mechanisms will account most

⁶ The secondary effect of differences in the trunk-to-head relation on the position of the eyes, which was demonstrated in monkeys by de Kleijn (1920), was demonstrated in human infants by Voss (1927).

economically for the phenomena which have been described in this paper. The sitting or standing posture of the average person functions like an "attitude" which has been imposed on the body by the head. The procedures employed in the experimental movements, by releasing the head from its habitual attitude, facilitate the righting reflexes and bring the subject into a different orientation within the gravitational field. The changes in breathing, in circulation, and in the use of the eyes, which are sometimes reported, take place automatically by reflex facilitation when the head moves into its new relationship to the trunk.

PSYCHOLOGICAL CONSIDERATIONS

I have described a mechanism by which the antigravity response can be facilitated. It is a mechanism which ordinarily operates below the level of consciousness. Magnus was emphatic that the righting reflexes are subcortical and inaccessible to direct conscious control.⁷ An indirect control can be

⁷ Magnus (1925):

It seems to be of the greatest importance that the whole central apparatus for the righting function (with the only exception of the optical righting reflexes) is placed subcortically in the brain stem and by this means withdrawn from voluntary action. The cortex cerebri evokes during ordinary life a succession of phasic movements, which tend to *disturb* the normal resting posture. The brain-stem centres will in the meantime *restore* the disturbance and bring the body back into the normal posture, so that the next cortical impulse will find the body prepared to start again [p. 349].

Magnus (1930):

We have . . . a subcortically acting apparatus which controls and adjusts the position of our body, whether erect or recumbent, in relation to space. This unconsciously acting mechanism, by the co-operation of complicated reflexes, restores our body to the normal position whenever it is displaced [p. 103].

established, however, if the subject learns to recognize and inhibit maladaptive postural sets which interfere with the response of the organism to gravity.

In the course of motor learning, sets may be developed which are not the best preparation for the movement to come. They may, in fact, hamper the execution of the movement. Unfortunately, like some of the "superstitious" responses which appear during conditioning experiments, such sets do not extinguish readily. The organism adapts to them quickly; they come to "feel right"; and they remain undetected, because once the stimulus to move has been received, attention becomes focused on the goal to be reached.

One of these inappropriate sets is the tendency to shorten certain neck muscles as a preparation for a movement against gravity. It has been demonstrated that such a movement is facilitated when the preparatory shortening of muscles is prevented by the experimenter. If the subject becomes aware of the tendency, he can learn to prevent it and thus establish an indirect control over the postural mechanism. In my experience, the only satisfactory way to achieve such a control is to reorganize the field of attention, so that when a stimulus to move is received, the focus of attention remains within the organism. This does not mean that the goal is excluded from attention; it means that the goal is not allowed to dominate the field. Attention is organized around the head-trunk relation, with extension in time and space so that both the stimulus and the response can be comprised within the same field.

Ordinarily, attention is directed either outward to the environment or inward to the organism itself. The central nervous system, however, is receiving information about movements

and positions of the body and its parts at the same time that it receives information about events in the world outside. There is no reason why the field of attention cannot be organized in the same way. In such a field, the relation of the head, neck, and trunk, kinesiologically perceived, forms the background against which events outside and inside the organism take place. Thus it is possible to perceive an object simultaneously with the organism's reaction to it, since both are comprised within the same field.

Perceptually, objects are known to vary with the psychological context in which they are perceived. A staircase, for example, is perceived differently depending on whether it is to be climbed or merely to be looked at. The difference, of course, lies within the perceiving organism. If it is a staircase to be climbed, it may elicit a postural set which is so marked that it can be detected by an outside observer. The set can be detected by the climber himself only if he reorganizes his field of attention so as to take in both the staircase and his reaction to it as he approaches and climbs it. With this shift in the focus of attention, he can perceive the cause-and-effect relation between the stimulus (the staircase) and his immediate response (the postural set). If he takes an experimental approach, he can devise a means to inhibit the set while continuing to make the specific response (climbing the stairs). In the process, the antigravity response will be facilitated in the same way as in the guided movements which were described above.

Climbing a staircase was used to illustrate a principle. Any activity—reaching for a pencil or making a speech—would have done as well. The principle of inhibition can be applied to any movement. A movement pat-

tern is a complex whole which, for convenience, may be thought of as having two aspects or parts: a specific, goal-directed part, and a tonic or postural part, by which the integrity of the organism is maintained while the specific response is being carried out. The tonic or postural part of a movement is not ordinarily perceived. Tensional patterns which interfere with the smooth performance of movement can be perceived, however. If they are inhibited, postural tonus is redistributed, and the specific, goal-directed response becomes easier. In contrast to the ease of the facilitated movement, old ways of moving come to feel wrong, and a new sensory standard is gradually established.

In the paradigm of postural change which I have just outlined, inhibition is the basic principle. Inhibition is a term which has been used by psychologists and physiologists in a variety of meanings (Diamond, Balvin, & Diamond, 1963). I have used it here to describe a process by which a person consciously refrains from making a response which he could make if he chose. In this sense inhibition is the central function of

a nervous system which, when it functions well, is able to exclude maladaptive conflict without suppressing spontaneity [Diamond et al., p. 395].

In the presence of inhibition, a stimulus should elicit only a generalized increase of alertness, leaving the organism free to respond or not respond.

The principle of inhibition, as it has been developed here, offers a new approach to the problem of behavioral change. In the close connection between inhibition and postural tonus is a mechanism which not only reveals the inner pattern of a stereotyped response but brings it under conscious control. In so doing, it greatly enlarges the

area of behavior where free choice can operate.

REFERENCES

- ALEXANDER, F. M. *Constructive conscious control of the individual*. New York: Dutton, 1923.
- ALEXANDER, F. M. *The use of the self*. New York: Dutton, 1932.
- ALEXANDER, F. M. *The universal constant in living*. New York: Dutton, 1941.
- BARLOW, W. Knowing how to stop. *Medical Press*, 1945, 214, 42-43.
- BARLOW, W. Anxiety and muscle tension. In *Modern trends in psychosomatic medicine*. London: Butterworth, 1954. Pp. 285-309.
- CARRINGTON, W. H. M. The F. Matthias Alexander technique. *Systematics*, 1963, 1, 233-247.
- CHATFIELD, P. O. *Fundamentals of clinical neurophysiology*. Springfield, Ill.: Charles C Thomas, 1957.
- COGHILL, G. E. Appreciation: The educational methods of F. Matthias Alexander. In F. M. Alexander, *The universal constant in living*. New York: Dutton, 1941. Pp. xxi-xxviii.
- DART, R. A. The attainment of poise. *South African Medical Journal*, 1947, 21, 74-91.
- DART, R. A. Voluntary musculature in the human body: The double spiral arrangement. *British Journal of Physical Medicine and Industrial Hygiene*, 1950, 13, 265-268.
- DE KLEIJN, A. On the effect of tonic labyrinthine and cervical reflexes upon the eye muscles. *Koninklyke Akademie van Wetenschappen, Amsterdam Proceedings*, 1920, 23, 509.
- DEWEY, J. *Human nature and conduct*. New York: Holt, 1922. Ch. 2.
- DEWEY, J. Introduction. In F. M. Alexander, *Constructive conscious control of the individual*. New York: Dutton, 1923. Pp. xxi-xxxiii.
- DEWEY, J. Introduction. In F. M. Alexander, *The use of the self*. New York: Dutton, 1932. Pp. xiii-xix.
- DIAMOND, S., BALVIN, R. S., & DIAMOND, F. R. *Inhibition and choice*. New York: Harper & Row, 1963.
- DUCHENNE, G. B. *Physiologie des mouvements*. (Orig. publ. 1867) [Physiology of motion.] (Trans. by E. B. Kaplan) Philadelphia: W. B. Saunders, 1959.
- FUKUDA, T. *Stato-kinetic reflexes in equilibrium and movement*. Tokyo: Igaku Shoin, 1957.
- GESELL, A. The ontogenesis of infant behavior. In L. Carmichael (Ed.), *Manual of child psychology*. New York: Wiley, 1954. Pp. 335-373.
- GRAY, H. *Anatomy of the human body*. (24th ed.) Philadelphia: Lea & Febiger, 1942.
- HELLEBRANDT, FRANCES A., SCHADE, MAJA, & CAIRNS, MARIE L. Methods of evoking tonic neck reflexes in normal human subjects. *American Journal of Physical Medicine*, 1962, 41, 90-139.
- HERRICK, C. J. *George Ellett Coghill*. Chicago: Univer. Chicago Press, 1949.
- HOWELLS, W. W. The designation of the principal anthropometric landmarks on the head and skull. *American Journal of Physical Anthropology*, 1937, 22, 477-494.
- HUXLEY, FRANCES M. On the reflex nature of apnoea in the duck in diving. *Quarterly Journal of Experimental Physiology*, 1913, 6, 147-196.
- IRVING, L. Respiration in diving mammals. *Physiological Review*, 1939, 19, 112-134.
- JONES, F. P. A mechanism for change. In P. Sorokin (Ed.), *Forms and techniques of altruistic and spiritual growth*. Boston: Beacon Press, 1954. Pp. 177-187.
- JONES, F. P. The influence of postural set on pattern of movement in man. *International Journal of Neurology*, 1963, 4, 60-71.
- JONES, F. P. Die kinästhetische Wahrnehmung von Haltung und Bewegung. In *Eutonie*. Ulm/Donau: Karl F. Haug, 1964. Pp. 100-124.
- JONES, F. P., & GILLEY, P. F. M. Head balance and sitting posture: An x-ray analysis. *Journal of Psychology*, 1960, 49, 289-293.
- JONES, F. P., GRAY, FLORENCE E., HANSON, J. A., & O'CONNELL, D. N. An experimental study of the effect of head balance on patterns of posture and movement in man. *Journal of Psychology*, 1959, 47, 247-258.
- JONES, F. P., GRAY, FLORENCE E., HANSON, J. A., & SHOOP, J. D. Neck muscle tension and the postural image. *Ergonomics*, 1961, 4, 133-142.
- JONES, F. P., & HANSON, J. A. Time-space pattern in a gross body movement. *Perceptual and Motor Skills*, 1961, 12, 35-41.
- JONES, F. P., & HANSON, J. A. Note on the persistence of pattern in a gross body movement. *Perceptual and Motor Skills*, 1962, 14, 230.
- JONES, F. P., HANSON, J. A., & GRAY, FLORENCE E. Head balance and sitting

- posture: II. The role of the sternomastoid muscle. *Journal of Psychology*, 1961, 52, 363-367.
- JONES, F. P., HANSON, J. A., & GRAY, FLORENCE E. Startle as a paradigm of malposture. *Perceptual and Motor Skills*, 1964, 19, 21-22.
- JONES, F. P., HANSON, J. A., MILLER, J. K., & BOSSOM, J. Quantitative analysis of abnormal movement: The sit-to-stand pattern. *American Journal of Physical Medicine*, 1963, 42, 208-218.
- JONES, F. P., & KENNEDY, J. L. An electromyographic technique for recording the startle pattern. *Journal of Psychology*, 1951, 32, 63-68.
- JONES, F. P., & NARVA, M. Interrupted light photography to record the effect of changes in the poise of the head upon patterns of movement and posture in man. *Journal of Psychology*, 1955, 40, 125-131.
- JONES, F. P., & O'CONNELL, D. N. Applications of multiple-image photography in the time-motion analysis of human movement, with a note on "color coding." *Photographic Science and Technique*, 1956, 3 (Ser. 2), 11-14.
- JONES, F. P., & O'CONNELL, D. N. Posture as a function of time. *Journal of Psychology*, 1958, 46, 287-294.
- KOPPÁNYI, T., & DOOLEY, M. S. Submergence and postural apnea in the muskrat. *American Journal of Physiology*, 1929, 88, 592-595.
- LANDAU, A. Über einen tonischen Lagereflex beim älteren Säugling. *Klinische Wochenschrift*, 1923, 2, 1253-1255.
- LANDIS, C., & HUNT, W. *The startle pattern*. New York: Farrar & Rinehart, 1939.
- MCCORMACK, E. E. Frederick Matthias Alexander and John Dewey: A neglected influence. Unpublished doctoral dissertation, University of Toronto, 1959.
- MCCOUCH, G. P., DEERING, I. D., & LING, T. H. Location of receptors for tonic neck reflexes. *Journal of Neurophysiology*, 1951, 14, 191-195.
- MACDONALD, P. Instinct and functioning in health and disease. *British Medical Journal*, 1926, 2, 1221-1223.
- MAGNUS, R. *Körperstellung*. Berlin: Springer, 1924.
- MAGNUS, R. Animal posture. *Proceedings of the Royal Society of London*, 1925, 98 (Ser. B), 339-353.
- MAGNUS, R. The physiological *a priori*. Lane lectures on experimental pharmacology and medicine: III. *Stanford University Publications*, 1930 (Series V. 2, No. 3), 331-337.
- MAGNUS, R., & DE KLEIJN, A. Die Abhängigkeit des Tonus der Extremitätenmuskeln von der Kopfstellung. *Pflügers Archiv für die gesamte Physiologie der Menschen und der Tiere*, 1912, 145, 455-548.
- PEIPER, A. Reflexes of position and movement. In *Cerebral function in infancy and childhood*. (Trans. by B. Nagler & Hilde Nagler) New York: Consultants Bureau, 1963. Ch. 4.
- RADEMAKER, G. G. J. *Das Stehen*. Berlin: Springer, 1931.
- ROMER, A. S. *The vertebrate body*. Philadelphia: W. B. Saunders, 1949.
- RUGG-GUNN, A. F. Matthias Alexander and the problem of animal behaviour. *Medical Press*, 1940, 203 (No. 5265), 285-287.
- SCHALTENBRAND, G. Normale Bewegungs- und Haltungs- und Lagereaktionen bei Kindern. *Deutsche Zeitschrift für Nervenheilkunde*, 1925, 87, 23-42.
- SCHOLANDER, P. F. The master switch of life. *Scientific American*, 1963, 209, 92-106.
- SHERRINGTON, C. S. *The endeavour of Jean Fernel*. London: Cambridge University Press, 1946.
- SIMONS, A. Kopfhaltung und Muskeltonus. *Zeitschrift für die gesamte Neurologie und Psychiatrie*, 1923, 80, 499-549.
- TOKIZANE, T., MURAO, M., OGATA, T., & KONDO, T. Electromyographic studies on tonic neck, lumbar and labyrinthine reflexes in normal persons. *Japanese Journal of Physiology*, 1951, 2, 130-146.
- Voss, O. Geburtstrauma und Gehörorgan. *Acta oto-laryngologica*, 1927, 11, 73-108.
- WALSHE, F. M. R. On certain tonic or postural reflexes in hemiplegia with special reference to the so-called associated movements. *Brain*, 1923, 46, 1-37.
- WELLS, H. S. The demonstration of tonic neck and labyrinthine reflexes and positive heliotropic responses in normal human subjects. *Science*, 1944, 99, 36-37.

(Received May 27, 1964)

COGNITIVE DEPENDENCE ON LINEAR AND NONLINEAR CUES¹

KENNETH R. HAMMOND AND DAVID A. SUMMERS

Institute of Behavioral Science, University of Colorado

Analysis of the cognitive process of inductive inference should focus on inferences drawn from nonlinear as well as linear relations. Brunswik's probabilistic functionalism is demonstrated as a conceptual and methodological framework within which this question can be investigated. Analysis of Ss' utilization of nonlinear relations is illustrated by studying 30 Ss in the following task: (a) one cue related in a linear, the other in a nonlinear manner to a criterion, (b) the criterion partly, but not perfectly, predictable from either cue alone, and (c) the criterion perfectly predictable from appropriate utilization of both. Results indicate that Ss can improve both overall performance and nonlinear data utilization, and that performance varied with task-relevant instructions.

This paper is concerned with the analysis of the cognitive process of inductive inference; it is focused upon the methodological and conceptual issues involved in the investigation of the process by which human subjects draw inferences from linear and nonlinear relations. The issues are discussed within the context of clinical inference and multiple-cue probability learning.

CLINICAL INFERENCE

Interest in the clinical versus statistical prediction issue has shifted somewhat from a concern with the relative *accuracy* of the two methods (Meehl, 1954; Sarbin, 1942) to a more detailed analysis of the *process* of clinical inference, or to what Meehl (1960) has called "The Cognitive

Activity of the Clinician." To what extent this process enables human subjects to utilize nonlinear as well as linear relations has become a focal point in this issue (cf. Meehl, 1960, p. 24). For although test manuals and clinical teaching emphasize the nonlinearity of certain cue-criterion relations,² there has been virtually no effort made to discover to what extent clinicians or other human subjects can effectively utilize such nonlinear relations in making inductive inferences. Indeed, studies of clinical inference strongly suggest that it is no more than a linear cognitive process. Thus, Todd (1954), Hammond (1955), Hoffman (1960), Grebstein (1963), Hursch, Hammond, and Hursch (1964), Hammond, Hursch, and Todd (1964) have found that linear (i.e., multiple regression) models are sufficient to account for the process. Also, Newton (in press) and Lee and Tucker (1962) have shown this to be the case in quasi-clinical tasks using college students as subjects. Anderson (1962)

¹ The research reported here was undertaken in the Behavior Research Laboratory, Institute of Behavioral Science, University of Colorado, and is Publication 51 of the Institute. This research was supported by Research Grant M-4977 from the National Institute of Mental Health. It is a part of the general program of research on inferential processes now in progress at the Behavior Research Laboratory. The authors are indebted to Frederick J. Todd, Cameron Peterson, and Richard Jessor for their assistance.

² Klopfer, Ainsworth, Klopfer, and Holt (1954), for example, suggest that a nonlinear relation exists between $W\%$ on the Rorschach and organizational effort when they indicate that both high and low values of $W\%$ reflect "little effort to organize experience [p. 300]."

likewise reports high predictability of a linear, additive model in impression-formation studies.

PROBABILITY LEARNING

Consideration of the results of multiple-cue probability studies leads to a similar conclusion; studies by Smedslund (1955), Summers (1962), Uhl (1963), Rappoport (1963), Todd and Hammond (in press), and Peterson, Hammond, and Summers (1964, in press) also show high linearity in the response system of subjects.

It must be emphasized, however, that in the above studies of clinical inference and learning the task itself has involved primarily *linear* relations; there were no compelling reasons for the subjects' response systems to be anything but linear. In view of Brunswik's (1956) admonition that tasks should be representative of a wide range of conditions, the performance of the subjects in situations involving nonlinear as well as linear relations should be investigated before concluding that the process of inductive inference is primarily linear. In short, having discovered what apparently is a strong tendency for human subjects to utilize the data from linear relations in a highly linear manner, it remains to investigate whether subjects utilize the data from nonlinear relations in a linear and/or nonlinear manner.

We need not restrict ourselves to the question of whether human subjects can utilize nonlinear as well as linear relations, however; we may go so far as to ask whether humans can combine *both* kinds of data and utilize both linear and nonlinear cognitive processes in the same task. That is, can human subjects combine data from linear and nonlinear cue-criterion relations concurrently in such a way as to make appropriate use of both? To the best of our knowl-

edge, this question has never been investigated empirically. It needs investigation, however; even brief reflection suggests that it is a question which lies at the heart of a large number, possibly a majority, of clinical inference tasks, as well as inference tasks in general. For if the clinician always combines data in a linear manner *irrespective of the form of the data*, the only difference in accuracy of prediction between the clinician and, say, a multiple regression equation would lie in differential approximation to optimum weighting of cues. Moreover, it would be unreasonable to assume that inferences required for adaptation to the human ecology neatly separate themselves into two mutually exclusive kinds—those with only linear relations and those with only nonlinear relations. A first step toward an empirical analysis of this problem is undertaken here.

CONCEPTUAL AND METHODOLOGICAL FRAMEWORK

Fundamental to the problem of investigating the question of linear and nonlinear inference is (a) the choice of a conceptual and methodological framework, and (b) a method of analysis. The purposes of this article are two: To demonstrate that Brunswik's (1956) theory of probabilistic functionalism makes it possible to investigate the question, and to show that a method of analysis developed within that framework makes it possible to increase our knowledge of subjects' abilities to cope with linear and nonlinear data. Our purposes, therefore, are mainly methodological.

Space does not permit a recapitulation of Brunswik's approach (1952, 1955, 1956; see also Hammond, 1955, for an application of Brunswik's theory and method to the problem

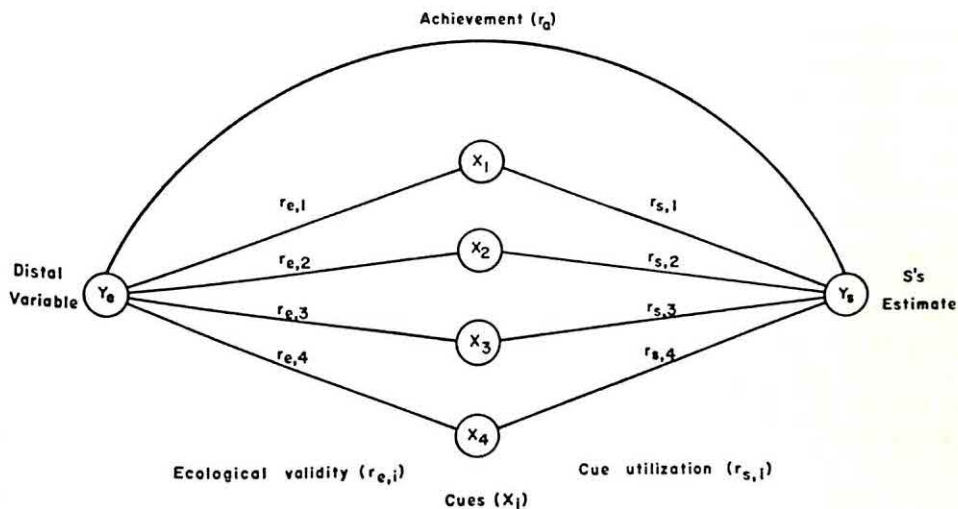


FIG. 1. Brunswik's lens model.

of clinical inference). Figure 1, however, shows how Grebstein's (1963) study of clinicians' efforts to predict IQ from Rorschach protocols may be incorporated in the Brunswikian "lens model." Thus r_e indicates the correlation between cues (e.g., $F + \%$) and the variable-to-be-estimated (IQ), and r_s indicates the correlation between cues and the clinician's estimate of IQ. Furthermore, a multiple correlation coefficient (R_e) may be calculated between cues and the criterion variable, and a multiple correlation coefficient (R_s) may be calculated between the cues and the clinician's estimates (Y_s) of the patient's IQ (Y_e). Finally, r_a indicates achievement, that is, the correlation between the clinician's judgment and the criterion values.

Given the above framework, Hursch et al. (1964) made a mathematical analysis of the limits of achievement under various conditions. Their analysis shows that achievement (r_a) is determined by the components included in the following equation:

$$r_a = \frac{R_e^2 + R_s^2 - \Sigma d}{2} + C\sqrt{(1 - R_e^2)(1 - R_s^2)} \quad [1]$$

where

r_a = the correlation between subject's judgment and the variable estimated

R_e = the multiple correlation between the cues and the variable estimated

R_s = the multiple correlation between the cues and subject's judgments

Σd = the sum of the products $(r_{e_i} - r_{s_i})(\beta_{e_i} - \beta_{s_i})$ where r_{e_i} = the correlation between cue_i and the variable estimated,

r_{s_i} = the correlation between cue_i and subject's judgment, β_{e_i} = the beta weight for the correlation between cue_i and the variable estimated, and β_{s_i} = the beta weight for the correlation between cue_i and subject's response

C = the correlation between the variance unaccounted for by the multiple correlation in the ecology and the variance unaccounted for by the multiple correlation in subject's response system

One further point needs to be made concerning the above equation. It should be noted that the first term on the right

$$\left[\frac{R_e^2 + R_s^2 - \Sigma d}{2} \right]$$

denotes the linear component of achievement (r_a). The second term on the right

$$[C\sqrt{(1 - R_e^2)(1 - R_s^2)}]$$

denotes the nonlinear component of achievement. The sum of the two

terms, of course, equals total achievement expressed as r_a . Thus the performance of subjects may be investigated to determine whether achievement is mainly a function of the linear component or the nonlinear component.³

When this method of analysis was applied to Grebstein's (1963) data, it was clear that because (a) R_e^2 was rather high (.79), (b) the average R_s^2 for the clinicians was also high (above .90 for all 15 clinicians studied), and (c) the Σd value was rather small, the clinicians could hardly have improved their performance by attempting to discover and utilize whatever nonlinear relations existed in the Rorschach IQ system (see Hammond et al., 1964, for complete details). When the above analysis was applied to Todd's clinicians, who were also attempting to predict IQ from the Rorschach, it showed that the achievement of these clinicians was *hampered* by attention to nonlinear characteristics of the task (see Hursch et al., 1964). In the above studies, however, response to nonlinear relations is a somewhat ephemeral notion. Its statistical denotation is clear, but its empirical meaning is not.

Statistically, it is the term C in the above equation which refers to nonlinearity. The C indicates the correlation between the residual variance in the clinician's response system and in the test-criterion system, that is, the relation between the variance not accounted for by the linear relations measured by R_e^2 and R_s^2 . Inasmuch as C reflects the appropriateness of the nonlinear or configural aspect of the clinician's response system, the value of this measure represents the only possible advantage the clinician may

have over the traditional multiple regression equation, which, of course, describes only linear relations.

Empirically, however, the measure (C) is yet unclear. In the Todd (1954) and Grebstein (1963) studies, for example, the value of C does not tell us *which* of the several possible nonlinear relations in the Rorschach IQ system were correlated with *which* nonlinear relations in a clinician's response system. Therefore, an experiment is called for in which it is not only possible to ascertain whether the subject can combine both linear and nonlinear data, but also one in which it will be possible to analyze and evaluate quantitatively the subject's relative dependence on both linear and nonlinear cue-criterion relations when they are under experimental control. Such an experiment and its results will be described below.

The present study deals with a very simple two-cue situation involving both linearity and nonlinearity. The task involves the following properties:

1. Criterion partly, but not perfectly, predictable from either predictor taken alone (cf. Horst, 1954; Meehl, 1950);
2. One predictor (X_1) linearly related to the criterion, one predictor (X_2) nonlinearly related to the criterion;
3. Criterion perfectly predictable if probabilistic cues combined according to the appropriate rule: $Y = X_1 + \sin X_2$;⁴
4. Fifty percent of the variance in the criterion determined by the linear predictor, 50% of the variance in the

⁴ $\sin X_2$ is an approximation to one phase of a sine curve. Using whole numbers and restricting the range of transformation to 10 digits, the correlation between X_2 and the criterion is not exactly .00. The contribution of this cue to the R_e^2 , however, is negligible: approximately .01.

³ See Tucker (1964) for a detailed discussion of the analysis of linear and nonlinear components of achievement.

criterion determined by the nonlinear predictor.

The first step was to ascertain the empirical validity of the components of the equation set forth above, with particular reference to C , the measure of nonlinear relations between the task system and the response system. For this purpose, an analysis of the performance of hypothetical subjects who pursue various types of response strategies was carried out by computer simulation (see Table 1).

Although it is doubtful that individuals ever follow specific inference strategies with complete consistency, these simulated strategies serve to delineate the components of a given strategy, as well as its predictive accuracy. Thus, perfect utilization of either the linear or the nonlinear cue alone will account for approximately 50% of the criterion variance; higher accuracy is possible only if the subject utilizes both types of relationships. For instance, an achievement (r_a) of .89 is possible even if the cues are weighted incorrectly, provided the linear and nonlinear specification is correct. Achievement of 1.00 is possible only if the statistical properties of the task are matched perfectly.

Having marked out some ideal types of performance and shown the

empirical meaning of the statistic C as well as other components of achievement, we now turn to the question of how the subjects will perform. In order to approximate roughly three possible approaches to clinical training, the subjects were studied with the following questions in mind:

1. Can subjects learn to predict the criterion in a task of this nature having no prior information regarding the properties of the task?

2. Can subjects learn to predict if given the "theoretical" structure (linearity and nonlinearity) of the task?

3. Can subjects learn to predict if they are instructed as to how the "theoretical" structure applies to the specific task?

PROCEDURE

Subjects were given a two-cue inference task having the statistical properties described above. Stimulus materials consisted of 100 $3\frac{1}{2} \times 6$ inch cards divided into 5 blocks of 20 cards each. Printed on every card were two 10-point "test scales," with the correct criterion value (hypothetical trait) on the back of each card. Instructions were focused upon providing different types of information as implied by the three learning conditions described above. Thus, Group I was told simply to make inferences on the basis of the two test scores and that accurate prediction was possible. The subjects in Group II were told that both linear and nonlinear relationships were involved and that the highest accuracy possible was contingent upon utilization of both. Examples of linear and nonlinear prediction were provided. The subjects in Group III were also provided with information regarding linear and nonlinear relationships; in addition, the linear and nonlinear cues were identified. All subjects observed the correct criterion value following each prediction. For half of the subjects in each group, the linear cue was on the right; for the other half, the linear cue was on the left.

RESULTS

Achievement. Figure 2 shows the mean achievement correlation for each

TABLE 1
SIMULATED RESPONSE STRATEGIES AND
STATISTICAL RESULTANTS (FOR THE
TASK $Y_e = X_1 + \sin X_2$, AND
 $r_{x_1} = .71$, $r_{x_2} = -.11$,
 $R_e^2 = .51$)

Strategies	Resultants				
	r_{x_1}	r_{x_2}	R_e^2	C	r_a
$Y_e = X_1$	1.00	.00	1.00	.00	.71
$Y_e = 3X_1 + X_2$.95	.32	1.00	.00	.64
$Y_e = X_1 + X_2$.71	.71	1.00	.00	.42
$Y_e = 3X_1 + \sin X_2$.95	-.05	.90	1.00	.89
$Y_e = X_1 + \sin X_2$.71	-.11	.51	1.00	1.00
$Y_e = X_1 + 3(\sin X_2)$.32	-.14	.12	1.00	.89
$Y_e = \sin X_2$.00	-.15	.02	1.00	.71

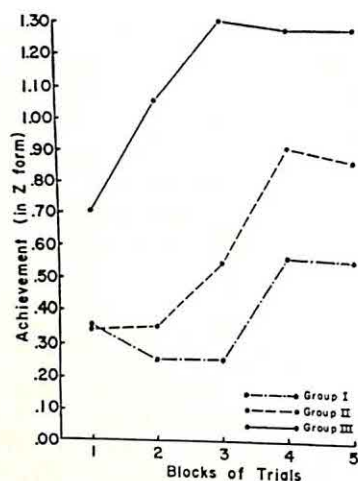


FIG. 2. Mean achievement correlations (in z form) for each condition over blocks of trials.

experimental group over the 5 blocks of 20 trials. Achievement was quite high for Group III ($\bar{x}_{ra} = .85$), somewhat lower for Group II ($\bar{x}_{ra} = .71$), and lowest for Group I ($\bar{x}_{ra} = .49$) on the final block of trials. Thus, as one would expect, the subjects given the most information (Group III) performed better ($p < .05$) than the subjects in the other two groups (see Table 2). There was an increase in

TABLE 2

ANALYSIS OF VARIANCE OF MEAN
ACHIEVEMENT CORRELATIONS
(IN z SCORES)

Source of variance	df	MS	F
Instruction conditions (A)	2	646.71	3.59*
Error (between)	27	179.84	
Blocks of trials (B)	4	104.51	5.88**
A \times B	8	21.88	1.23
Error (within)	108	17.75	

* $p < .05$.

** $p < .01$.

performance over all groups ($p < .01$), although Group I contributed least to this increase. There was no indication that one group learned more rapidly than another, since the interaction between groups and blocks was not statistically significant. The curves in Figure 2 suggest that learning had reached an asymptote by the end of the 100 trials, with a clear separation of all three groups.

Differential Cue Dependency. The extent to which the subjects utilized the linear (X_1) and nonlinear ($\sin X_2$) cues is illustrated in Figure 3. A Wilcoxon signed-ranks test indicates that the increase in the dependence on the linear cue from Block I to Block V

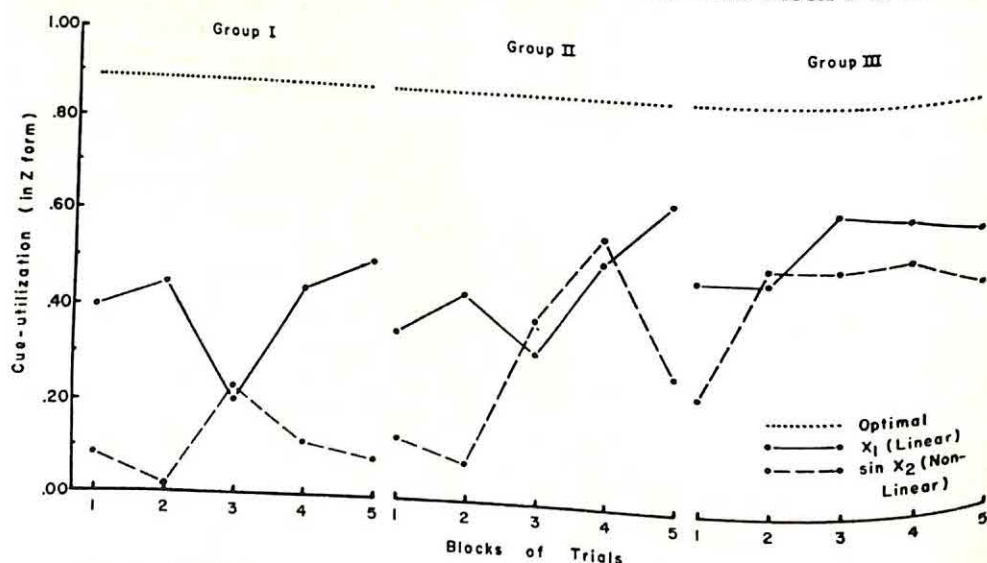


FIG. 3. Utilization coefficients for linear cue (X_1) and nonlinear cue ($\sin X_2$).

is significant ($p < .02$) only in Group II, and the increase in dependence upon the *nonlinear* cue from Block I to Block V is significant ($p < .01$) only in Group III. Thus, the nature of the information given the subject determines whether the linear or nonlinear cue will be *increasingly* utilized during learning. The graph in Figure 3 also indicates that there is a much stronger dependence on the linear cue than on the nonlinear cue in the minimum information condition (Group I); the curves are more widely separated here than in any other condition.

Note also that the utilization of both cues does not reach optimality. Perfect achievement would require that both curves reach the $z = .89$ ($r = .71$) line at the top of the graph. Overall, there appears to be a tendency to depend more on the linear cue.

Total Linearity. Figure 4 shows the mean linearity (R_s^2) of the subjects in each group. Clearly the different instructions did not differentially affect the extent to which the subjects combined the cues in a linear manner. Nor was the increase in R_s^2 from Block I to Block V statistically significant for any group. Figure 3 indicates

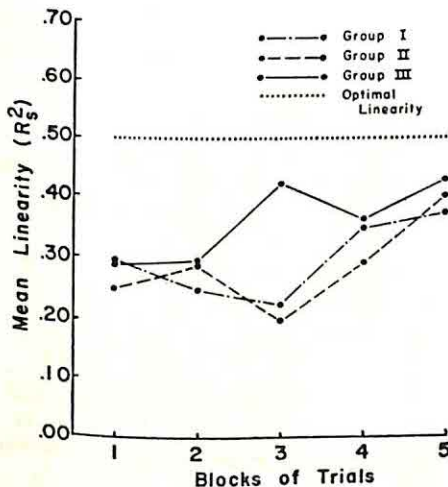


FIG. 4. Mean response system linearity for each condition over blocks of trials.

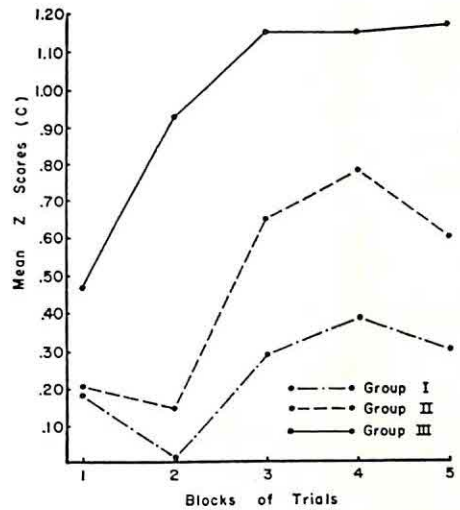


FIG. 5. Mean C correlations (in z form) for each condition over blocks of trials.

that in Group I a greater proportion of the total linearity (R_s^2) may be attributed to *linear* utilization of the *nonlinear* cue than in either of the two other groups.

Most important is the fact that the linearity of the subjects is less than optimal, and the absolute value of the mean R_s^2 is quite low.

Accuracy of Nonlinearity (C). Figure 5 illustrates the differential response to the nonlinear aspects of the task in the three groups. Group III clearly developed the largest value of C. The difference between the effect of instructions was statistically significant ($p < .05$; see Table 3). Trials also had an effect on the extent to which the subjects *correctly* utilized

TABLE 3
ANALYSIS OF VARIANCE OF MEAN C
CORRELATIONS (IN z SCORES)

Source of variance	df	MS	F
Instruction conditions (A)	2	726.95	4.20*
Error (between)	27	172.82	
Blocks of trials (B)	4	138.04	6.71**
A \times B	8	21.54	1.04
Error (within)	108	20.55	

* $p < .05$.

** $p < .01$.

the nonlinear cue, there being a significant ($p < .05$) increase in the value of C over all groups.

Linear and Nonlinear Components of Achievement. Figure 6 shows the relative contribution of the linear

$$\left[\frac{R_e^2 + R_s^2 - \Sigma d}{2} \right]$$

and nonlinear

$$[C\sqrt{(1 - R_e^2)(1 - R_s^2)}]$$

components of achievement (r_a) for each group.⁵ Appropriate weighting of the cues in this task would require that each component contribute equally to achievement. Inspection of these curves indicates that Groups II and III are clearly superior to Group I in adapting to the equal contributions of the two components of the task.

DISCUSSION

The issue of additivity versus pattern plays a central role throughout the history of psychology. Because psychologists have persistently taken an either-or posture toward this matter, it was only natural for them to take an either-or posture toward the linearity versus nonlinearity issue, as they have toward the clinical versus statistical prediction issue. All three are merely different versions of the same problem. Yet psychological experiments have shown that humans can solve problems of inductive inference in *either* an additive (linear) *or* patterned (nonlinear) manner.

It is possible to demonstrate pure cases of either process, however, only in very peculiar circumstances such

⁵ In order to express, in z score form, the two components in such a way that they add to produce r_a , also in z score, a transformation to z for each component was based on the proportion each contributed to r_a in raw score form.

as psychological experiments. While laboratory demonstrations of either process in pure form may be of some interest, the very peculiarity of the circumstances necessary to elicit them makes their relevance dubious. Those who wish to investigate the human capacity for inductive inference will not wish to restrict themselves to conditions which evoke *either* one process *or* the other. More to the point, we believe, is the analysis of inference in situations where *both* linear and nonlinear cognitive processes can contribute to problem solution. What is needed is the investigation of the human's capacity to bring both processes to bear on problems which require both for appropriate functioning.

The specific results of the present experiment indicate that achievement can be quite high in a task of this sort, but that achievement is contingent upon information about the task. Such information also determines whether the linear cue or the nonlinear cue will be increasingly utilized during learning.

The mean linearity (R_s^2) of the subjects was less than optimal and quite low, ranging from .21-.42. This finding is important because although results of previous studies indicate a high degree of linearity on the part of the subjects in multiple-cue probability learning tasks, the present results indicate that the propensity for a highly linear, additive response system is contingent upon the subject being presented with a highly linear task system.

The results support the empirical significance of C (see Equation 1) as a measure of the relation between nonlinear variances in the task and response systems of the subject. The C increases as a function of both instructions and trials in a situation in which the nonlinear cue is known and

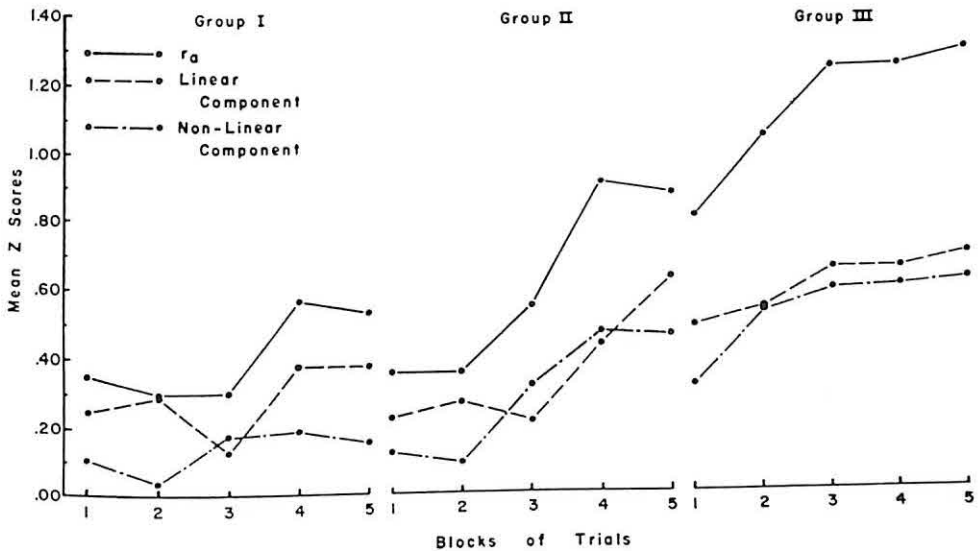


FIG. 6. Achievement correlations and component scores for each condition over blocks of trials.

specified by the experimenter, a result not previously available.

Finally, the results indicate that the relative contributions to accuracy of the linear and nonlinear components of performance were about equal, except in the minimal information condition, where the main contribution to achievement came from the linear component. In view of the fact that the linear and nonlinear components of the task itself were equally balanced, this result provides unexpected and impressive evidence of cognitive adjustment.

Moreover, results of current studies (Summers & Hammond, 1964) employing tasks with an *imbalance* between linear and nonlinear components also point toward matching of the linear and nonlinear components of achievement with the respective components of the task.

Generalization of all the above specific findings must, of course, be restricted to the specific conditions of the task; only one linear and one nonlinear cue was present, and the non-

linear function chosen is only one of a multitude of nonlinear functions confronting the human problem solver. (See Hoffman, 1960; Meehl, 1954, for discussion of possible nonlinear relations representative of clinical inference.) In addition to number of cues and type of nonlinear function, type of feedback (statistical as opposed to the traditional outcome feedback employed here; see Newton, in press; Todd & Hammond, in press) and amount of linearity (R_e^2 ; see Hursch et al., 1964) are certainly relevant dimensions along which investigation should be focused. It has been the purpose of this paper to demonstrate that such investigations are possible within the conceptual and methodological framework of Brunswik's probabilistic functionalism and that the necessary analytical tools are available.⁶

⁶ A (Fortran) computer program for calculating the statistics discussed in this article is available upon request at the Institute of Behavioral Science, University of Colorado.

REFERENCES

- ANDERSON, N. Application of an additive model to impression formation. *Science*, 1962, 138, 817-818.
- BRUNSWIK, E. *The conceptual framework of psychology*. Chicago: Univer. Chicago Press, 1952.
- BRUNSWIK, E. Representative design and probabilistic theory. *Psychological Review*, 1955, 62, 236-242.
- BRUNSWIK, E. *Perception and the representative design of psychological experiments*. Berkeley: Univer. California Press, 1956.
- GREBSTEIN, L. Relative accuracy of actuarial prediction, experienced clinicians, and graduate students in a clinical judgment task. *Journal of Consulting Psychology*, 1963, 37, 127-132.
- HAMMOND, K. R. Probabilistic functioning and the clinical method. *Psychological Review*, 1955, 62, 255-262.
- HAMMOND, K. R., HURSCH, CAROLYN J., & TODD, F. J. Analyzing the components of clinical inference. *Psychological Review*, 1964, 71, 438-456.
- HOFFMAN, P. The paramorphic representation of clinical judgment. *Psychological Bulletin*, 1960, 57, 116-131.
- HORST, P. Pattern analysis and configural scoring. *Journal of Clinical Psychology*, 1954, 10, 3-11.
- HURSCH, CAROLYN J., HAMMOND, K. R., & HURSCH, J. Some methodological considerations in multiple-cue probability studies. *Psychological Review*, 1964, 71, 42-60.
- KLOFFER, B., AINSWORTH, M., KLOFFER, W., & HOLT, R. *Developments in the Rorschach technique*. New York: World Book, 1954.
- LEE, JOAN C., & TUCKER, R. An investigation of clinical judgment: A study in method. *Journal of Abnormal and Social Psychology*, 1962, 64, 272-280.
- MEEHL, P. E. Configural scoring. *Journal of Consulting Psychology*, 1950, 14, 165-171.
- MEEHL, P. E. *Clinical versus statistical prediction*. Minneapolis: Univer. Minnesota Press, 1954.
- MEEHL, P. E. The cognitive activity of the clinician. *American Psychologist*, 1960, 15, 19-27.
- NEWTON, J. Judgment and feedback in a quasi-clinical situation. *Journal of Personality and Social Psychology*, in press.
- PETERSON, C., HAMMOND, K. R., & SUMMERS, D. Multiple probability learning with shifting cue weights. Behavior Research Laboratory Report, No. 54, University of Colorado, 1964. (Mimeo)
- PETERSON, C., HAMMOND, K. R., & SUMMERS, D. Optimal responding in multiple-cue probability learning. *Journal of Experimental Psychology*, in press.
- RAPPOPORT, L. Interpersonal conflict in a probabilistic situation. Unpublished doctoral dissertation, University of Colorado, 1963.
- SARBIN, T. R. A contribution to the study of actuarial and individual methods of prediction. *American Journal of Sociology*, 1942, 48, 593-602.
- SMEDSLUND, J. *Multiple probability learning*. Oslo, Norway: Oslo Univer. Press, 1955.
- SUMMERS, D., & HAMMOND, K. R. Inference under varying levels of linear and non-linear task variance. Behavior Research Laboratory Report, No. 56, University of Colorado, 1964. (Mimeo)
- SUMMERS, S. The learning of responses to multiple weighted cues. *Journal of Experimental Psychology*, 1962, 64, 29-34.
- TODD, F. A methodological analysis of clinical judgment. Unpublished doctoral dissertation, University of Colorado, 1954.
- TODD, F., & HAMMOND, K. R. Differential feedback in two multiple-cue probability learning tasks. *Behavioral Science*, in press.
- TUCKER, L. R. A suggested alternative formulation in the developments by Hursch, Hammond, and Hursch, and by Hammond, Hursch, and Todd. *Psychological Review*, 1964, 71, 528-530.
- UHL, C. N. Learning interval concepts: I. Effects of differences in stimulus weights. *Journal of Experimental Psychology*, 1963, 66, 264-273.

(Received June 9, 1964)

NEURAL CONSOLIDATION AND ELECTROCONVULSIVE SHOCK¹

DONALD J. LEWIS

AND

BRENDAN A. MAHER

Rutgers University

University of Wisconsin

It is believed by many that the neural engram undergoes a consolidated phase after a learning trial. During this phase the engram is particularly susceptible to disruption by traumatic incidents. A review of the literature on electroconvulsive shock (ECS) provides little, if any, support for this point of view. ECS does, however, result in retrograde amnesia (RA). If the RA is not due to disruption of a consolidation process, then what are its sources? The writers interpret the literature to mean the ECS produces an inhibition of the Pavlovian variety which follows the known laws of learning and becomes conditioned to stimuli of the ECS situation. This point of view is shown adequate to explain much of the data produced by ECS studies and is consistent with other interpretations of the effects of massive stimulus discharges.

Very few psychologists of any theoretical persuasion would deny that a learning trial produces some kind of change in the nervous system. Most would agree that this change must be a structural one to account for the relatively permanent modifications of behavior that constitute learning. However, there has been considerable difference of opinion about the nature of this neural residue of experience—the engram. Most physiological psychologists and those physiologists and biochemists who have shown an interest in learning seem to prefer an engram consolidation interpretation (Glickman, 1961). The consolidation interpretation is difficult to describe, for there are several versions of it, and none of them has been elaborated with any great degree of specificity.

All consolidation theorists seem to agree that immediately following a learning trial the engram is in a relatively fragile condition and is therefore susceptible to impairment or eradication. If it is impaired during this

period, memory loss or forgetting occurs. In a short period of time following the learning trial, however, the engram consolidates, becoming a relatively permanent residue of the learning experience and resistant to destruction. We need not at this moment be concerned whether the engram is molecular in nature, or cellular, electrical, or chemical, or any of the other many possibilities. The major notion that concerns us here is that the engram has an initial phase in which it is vulnerable and easily destroyed.

The consolidation theory of learning was first proposed by Müller and Pilzecker (1900) in order to account for the forgetting of verbal materials. They knew that a list of verbal materials interpolated between the learning of an original list and its recall would result in interference with the retention of the original list—the phenomenon of retroactive inhibition. Müller and Pilzecker hypothesized that the engrams produced by the recitation of the items from the original list were not yet consolidated and thus were susceptible to disruption produced by re-

¹ Supported in part by Grant MH-07129-01 from the National Institute of Mental Health.

citing items from the interpolated list. Although widely accepted for some years by experimenters on verbal learning, the consolidation theory has few adherents among this group today. Too few of the variables manipulated by the verbal learners show the temporal effects demanded by the consolidation theory.

The phenomenon of proactive inhibition, for example, is difficult to encompass within a consolidation framework. Retention of verbal units has been shown to depend upon the nature of the list of items previously memorized (Underwood, 1957). Here we have a memory deficit brought about by preceding events—by events occurring before the second-list engram was in existence. Also the degree of similarity among lists of items has been shown to have a great effect upon retention. For these and other reasons, many workers in verbal learning (e.g., Postman, 1961; Underwood & Postman, 1960) espouse a competition theory of forgetting. The competition theory maintains that the learning of one response does not destroy or obliterate another. The “unlearning” that is said to take place is temporary and transitory (Postman, 1961) as indicated by the spontaneous recovery of what was unlearned and by the savings with which the unlearned may be relearned. The possible degree of consolidation of a neural engram is not involved in such a conceptualization.

One of the most important considerations in evaluating a consolidation hypothesis is the length of time during which consolidation may be presumed to take place. Many authorities believe that this time period is quite short. Pearlman, Sharpless, and Jarvik (1961), for example, say, “It is agreed that the interval during which such agents can exert significant retrograde effects is relatively short, a few

hours at most [p. 109].” Deutsch (1962) places it at about 1 hour while John (1961) reduces it to between $\frac{1}{2}$ –1 hour. Most experimental tests of the consolidation hypothesis have involved the manipulation of variables intended to speed up or slow down the fixation of the engram. The theory demands that these variables must have a greater effect when manipulated soon after a learning experience than when this is done some time later. Variables administered a considerable time after a learning experience—say 12 or more hours—but which nevertheless result in a memory deficit cannot be considered as producing data germane to a consolidation notion.

Glickman (1961) recently evaluated the evidence relevant to the consolidation theory, and it was his opinion that “experimentally induced RA (retrograde amnesia) has produced the best evidence for the existence of a consolidation process since the results would be predictable from perseveration theory, while the primary competing theory, the associative interference theory, has no explanation to offer [p. 220].” Glickman states that the majority of the evidence supporting a consolidation hypothesis comes from studies in which electroconvulsive shock (ECS) is used to produce RA, and it is this body of data that will be our main concern here. Our approach will be to take up the major experimental variables in which ECS has been involved and to evaluate the data derived from their manipulation insofar as it pertains to consolidation. Are there adequate data to substantiate the notion that a period of time following learning is necessary for an engram to become fixed?

Without adopting a consolidation approach, one can still inquire into the physiological processes underlying the phenomena of proactive inhibition.

Certainly physiological events do take place and, whatever they are, they must take place in time. The consolidation hypothesis is distinguished by the notion that the learning engram requires a period of time to become fixed and that this time period is long enough to permit external events to interfere physically with fixation, but still not so long as to endure almost indefinitely.

THE LEARNING-ECS INTERVAL

The interval between the end of a learning trial and the ECS is a crucial one because it is during this period that the engram is presumed to consolidate. Duncan (1949) manipulated this variable in a classic ECS experiment that gave support to a consolidation theory. Using a shuttle box, he gave his subjects one trial a day followed at different time intervals by ECS. The intervals between the learning trial and the ECS for the various groups were: 20 seconds, 40 seconds, 4 minutes, 15 minutes, 1 hour, 4 hours, and 14 hours. He found that the relationship between the number of successful avoidance responses and the time intervals was sharply negatively accelerated. There was no evidence of learning with the shortest interval, indicating total RA, but if 1 hour or more elapsed between the end of a trial and the convulsion, there was no apparent memory loss.

Gerard (1955) administered ECS at intervals of 5 minutes, 15 minutes, 1 hour, and 4 hours following maze learning. He found no learning at the shortest interval and no impairment at the longest. Thompson and Dean (1955) used a choice situation and gave a single ECS either 10 seconds, 2 minutes, 1 hour, or 4 hours upon completion of acquisition. They found no learning at the shortest interval and no impairment at the longest.

Heriot and Coleman (1962) used a one-trial avoidance situation in which the subjects were taught to avoid a bar which they had previously learned to press. One ECS followed the avoidance experience at either 1, 7, 26, 60, or 180 minutes. The three shortest intervals obliterated the effects of the avoidance training, the longest interval had no effect, and the 60-minute interval was intermediate. These studies, a representative sample of a number of others, are all consistent in showing retention to be a negatively accelerated function of the interval between learning and the administration of ECS. None of them indicates that ECS is effective with an interval as long as 4 hours, and they all can be considered to fulfill the assumptions which support a consolidation hypothesis, although there are other equally tenable explanations of these data.

There are a number of studies, however, which are not supportive of a consolidation hypothesis in that they have involved the administration of ECS at greatly extended intervals following the last learning trials. Most of these also show that ECS results in some impairment of a previously learned response. Carson (1957), for example, gave a series of ECSs, two a day over a 20-day period, which began 24 hours following the last learning trial. A marked deficit due to ECS appeared, and the deficit lasted over a period of 15 days at which time the experiment ended. Williams (1963) gave ECS to one group of subjects 24 hours after the completion of training, and to a second group 13 days later. The response, a conditioned emotional response (CER), was greatly attenuated by the ECS following training by one day and was still *slightly effective* by the ECS given 13 days later.

Geller, Sidman, and Brady (1955) allowed 3 days to intervene between

training and ECS and found that the learned response, a CER, nevertheless had disappeared as a result of the convulsive treatment. Very significantly the CER had returned when the subjects were retested 30 days later. The attenuating effect in this case was not the permanent one that should have resulted if an incompletely consolidated engram had been destroyed. Previously Brady (1951) had given ECS 30, 60, and 90 days following the learning of a CER. He found no immediate effect on the retention of the CER, although the three groups did align themselves regularly on a resistance to extinction response measure which would indicate some attenuating effect for the very long intervals used. Stone and Bakhtiari (1956) gave 40 acquisition trials on a 13-choice multiple-T water maze with a 24-hour inter-trial interval. When ECS was administered 30 days following the last acquisition trial (70 days after the first trial), a deficit appeared temporarily on a subsequent retention test, but the deficit disappeared with further training.

Braun, Patton, and Barnes (1952), using monkeys, gave 515 object-quality discrimination problems on the Wisconsin General Test Apparatus which took approximately 56 days to complete. By the end of this experience the subjects had developed a highly efficient learning set and could learn a new problem much quicker than they did at the beginning of training. Seven days after the training was over they were given a series of 20 convulsions, 3 a week. When retested 3 days after the last convulsion, the treated animals were markedly inferior to controls, and this inferiority lasted over a series of 96 more problems given 8 a day. In this experiment ECS was given 63 days after the completion of all learning trials. This is an enor-

mously long time for consolidation to continue.

We have seen that ECS interferes with a previously learned response when it is administered as long as 30 or even 60 days following the conclusion of learning. We have also seen that the deficit brought about by ECS can also be somewhat temporary. These two facts make difficulties for a consolidation theory. They indicate that retrograde amnesia can be brought about by ECS even when the engram associated with original learning has completely consolidated—assuming that engrams do consolidate. We will return to this point later.

ECS—LEARNING INTERVAL

The data of proactive inhibition were a source of great difficulty for those holding to a consolidation point of view as an explanation for the forgetting of verbal materials. The same difficulty seems to prevail using ECS. Poschel (1957) first gave his animals approach training in a runway. Then they received a series of 10 ECSs over 2 days. Next they were placed in the end box of the runway and given a nonconvulsive shock to the feet through a grid floor. Seven days later tests were made to determine if they would tend to avoid the end box where they had received foot shock. The animals who had received ECS showed no avoidance of the end box although an avoidance was clearly present in control subjects who did not receive ECS. The ECS, given *before* the punishment in the end box, interfered with the learning of the avoidance response.

Adams and Lewis (1962a) used a two-compartment shuttle box in which animals learned to run from the first compartment into the second because of a grid shock to the feet in the first. Before this shuttle training, one group of subjects was given a series of ECSs

in the first compartment of the shuttle box. Then, 3 days *later*, they were given avoidance trials. The animals given ECS had much greater difficulty learning the avoidance response than did another group which had not received ECS first. How can an engram be interfered with 3 days before it occurs? There is little possibility that the response interference could be due to the postconvulsive confusion reported for human patients since such confusion lasts very few hours at most (Krains, 1957, p. 466).

Leukel (1957) used a 14-unit multiple-T water maze in which rats received one trial a day for 10 days. Either 1 minute, 5 minutes, 30 minutes, or 2 hours after a daily learning trial they were given ECS. At the conclusion of this 10-day acquisition ECS phase the subjects were given further trials in the water maze, one a day, until they reached a criterion of three successive errorless trials. There followed a 30-day idle period and then more trials to same criterion in order to measure retention. The initial acquisition data showed a response depression for all groups independent of the trial-ECS interval, a finding not helpful to the consolidation notion. A more interesting result of this study for our present purposes was that the response depression resulting from ECS remained through both the initial relearning trials and the 10 retention (and relearning) trials given 30 days later. This is further evidence that the effects of ECS can interfere with later learning, i.e., proactively; in this case the effect of ECS remained and interfered with relearning as much as 30 days later.

SPACING OF ECS TRIALS

The prediction of a consolidation theory regarding the spacing of ECS experiences seems clear enough. The

closer ECS follows a learning trial and the more compact the series of convulsions, the greater should be the interference of ECS with an engram. An experiment by Brady, Hunt, and Geller (1954) brings some information to bear on this question.

After their animals had acquired a bar-press response they were presented with a clicking noise which was followed by shock to the feet. After a few trials of this sort, the subjects stopped pressing the bar (CER or conditioned emotional response) whenever the clicking occurred, although they would continue to press otherwise. Twenty-four hours after the CER was well developed, all groups of subjects received 21 ECSs, but with a different time interval between treatments for each group. One group received ECS every second, and other groups were convulsed at intervals of 30 minutes, 1, 8, 24, or 72 hours. The very interesting results showed virtually complete attenuation of the CER in the groups receiving ECS at 1-, 8-, or 24-hour intervals. ECS given every second, at 30-minute or 72-hour intervals did not produce as much effect. Some very complicated assumptions could be made about the course of consolidation in order to preserve the theory, but they would be so specific to this experiment that their value would be doubtful.

SPACING OF ORIGINAL LEARNING TRIALS

Thompson and Pennington (1957) gave intertrial acquisition intervals of 45 seconds, 2, 3, 4, 5, and 6 minutes on a visual discrimination problem and continued learning trials until an acquisition criterion was attained. ECS followed immediately. Their results indicated that ECS had less effect with the longer intertrial intervals for original learning. The experimenters at-

tributed their results to the increased time within the widely interspersed acquisition intervals permitting neural consolidation to occur. They say, "This follows from the theory if it is assumed that the amount of consolidation of the memory trace increases as a function of the duration of perseverative activity. Thus, a long intertrial interval allows the posttrial organic events to produce greater consolidation than a short intertrial interval [p. 401]."

An implication of this argument is that engrams even from identical training trials interfere with each other when they follow closely one after another. The evidence, however, is quite strong (e.g., Kimble, 1961) that the effect of massed acquisition trials is a temporary one and therefore due to performance variables rather than learning ones. This militates against a consolidation interpretation.

EXTINCTION OF RA

Adams and Lewis (1962b) showed very clearly that RA due to ECS need not be permanent. They showed, further, that the RA could be experimentally removed. Two groups of animals were given a series of acquisition trials in a shuttle box with each trial followed immediately by a convulsion. The experimental subjects were then returned, two at a time, to the start compartment of the shuttle box and left there for 5 minutes on 5 successive days. The other group remained during this time in their living cages. Then both groups were given reacquisition trials. The animals that had had extinction experience in the shuttle box relearned much faster than the controls. This indicates that the RA produced by ECS is not permanent, and that recovery can be controlled.

EMOTIONAL RESPONDING

There have been a host of studies which seem to indicate that ECS has a special antagonism for emotional responding. The first clear demonstration of this was made by Hunt and Brady (1951), and their series of researches in this area has contributed considerable information relative to consolidation theory. Brady and his colleagues used approximately the same experimental arrangement, which has already been mentioned briefly, for most of their convulsive studies. The subjects were first taught to press a bar on a variable-interval schedule. Then they were presented with a brief tone associated with a shock through a grid floor. After a few trials of this sort, the animals stopped responding as soon as the tone sounded and returned to responding upon the cessation of the tone. The response measured in these studies is a conditioned suppression (nonresponding) which Brady labels a conditioned emotional response (CER). Three or 4 days after the last conditioning trial, and when the CER is strong, the subjects are given a series of ECSs. Brady and his colleagues typically administer a series of 21 ECSs at the rate of 3 a day for 7 days. Then, typically, 2 days after the last ECS, the subject is returned to the Skinner box and the conditioned stimulus (CS) presented. The experimental animals, those that received the ECS, continue pressing at about the same rate as they did before the conditioning of the CER. The control animals stop pressing immediately upon presentation of the CS. These findings indicate that ECS serves to attenuate a CER but that it has no discernible effect upon the bar press. It would seem either that emotional responses are particularly sensitive to the effects of ECS, or that the bar press is not.

There are at least two features of the Brady et al. (1954) studies having unfavorable implications for a consolidation theory. First of all, there is the 3- or 4-day interval between the end of conditioning and the first administration of ECS. This is a long time for a neural trace still to be consolidating. Second, ECS has no clear effect on the bar press under Brady's conditions.

Another interesting experiment by Brady (1952), using the standard paradigm described above, shows that the CER actually recovers to some extent within 30 days after the cessation of ECS experience. This can only mean that the effects of ECS, in this situation, are transitory. Brady has also shown, as we have seen, that when ECS is given 30, 60, or 90 days following the conditioning of a CER, the CER shows no attenuation on a retention test given 4 days later. When the CER was subsequently extinguished, however, the rate of extinction was proportional to the learning ECS interval; the 30-day subjects extinguished fastest followed by the 60- and 90-day subjects in that order. Again the learning-ECS interval is important even though the shortest interval is 30 days.

Hunt, Jernberg, and Brady (1952) have shown that if extinction is given following ECS the CER will not recover, although the CER will recover following ECS if no extinction period is added. A careful look at the "close grain" of responding in the Hunt et al. (1952) study revealed that the CER was never totally obliterated by the ECS but was only greatly attenuated. Probably extinction had its effect upon this weakened CER. This study indicates that a process similar to spontaneous recovery may be applicable to the attenuated CER, and perhaps it also points to the similarity of the effects of ECS and extinction.

Thus, Brady's data show that when a CER is superimposed upon an instrumental response the ECS can attenuate the CER, leaving the instrumental response undisturbed. This finding has been duplicated by a host of investigators using quite a large number of experimental situations, and there seems to be absolutely no question about its reality. Heistad (1955), for example, trained animals to tilt a box which was balanced on a fulcrum. Tilting the box allowed water to flow into a trough and the thirsty animals could then drink. After the instrumental response was well acquired, a CER was instilled by means of shock through a grid floor in the apparatus; the animal now crouched and huddled instead of tilting the cage for water. Two days after the last CER trial the experimental animals were given a series of ECSs. When put back in the apparatus, the animals returned immediately to tilting the cage for water; the CER had been almost completely obliterated.

Williams (1961) first taught animals to approach food in a runway. Then they were shocked through a grid in the floor of the last section of the runway to produce a conflict demonstrated by "hesitations" as the animals approached this section. A series of ECSs commenced 1 day after the last conflict trial. The effect of the ECS was virtually to obliterate the hesitation without interfering with the overall instrumental running response. In another study, Heriot and Coleman (1962) trained subjects to press a bar in a Skinner box. Then they were given a single strong shock upon touching the bar which generated avoidance of the bar. ECS was given a short time following the bar shock, and when the animals were returned to the box they began to press the bar again; avoidance was gone but not the bar

press. Madsen and McGaugh (1961), in an experiment similar to that of Pearlman et al. (1961), placed their subjects on a small platform suspended over a charged metal plate. When the animals stepped off the platform, they received a shock to the feet through the metal plate. Their response was to withdraw from the edge of the suspended platform and not to step off it. They were then given a single ECS and returned to the platform upon recovery. Now they stepped off the platform again as if they had never received shock through the metal plate; again the CER was selectively affected by ECS.

There are several other experiments which are instructive concerning the selective effect of ECS; in these studies the instrumental response is also attenuated by ECS. Heistad (1955) trained animals to move from one compartment to another in a circular maze when a CS was presented. If they moved soon enough after the appearance of the CS, they avoided a shock delivered through the grid floor. After the avoidance response was well learned, the animals were given a series of ECSs which effectively reduced the avoidance response. In this study the active muscular responding—the avoidance—was attenuated by ECS. This is in contrast to the Brady type of study in which the superimposed CER is attenuated but the active bar press is not. Carson (1957) in a similar fashion attenuated a wheel-turning avoidance response by ECS, and in a host of other studies (e.g., Duncan, 1949) an active avoidance response has been at least depressed by ECS. It may be that both the instrumental response and the emotional response have been interfered with in these studies, although it is also possible that the reason that the instrumental avoidance response is reduced is because the fear

lying behind it has been attenuated by the ECS.

Anesthetics

The use of anesthetics along with ECS has produced some interesting information because some anesthetics can prevent convulsive behavior from occurring even though a considerable "convulsive" current is used.

Porter and Stone (1947) had their animals learn a multiple-U maze and then subjected them to ECS. One group experienced ECS while under ether anesthetic, and the other group was unanesthetized. The results showed that ECS had little effect upon retention when given under anesthetic, but had a great effect otherwise. Hunt, Jernberg, and Otis (1953) found that convulsions produced by carbon disulfide attenuated a CER in a normal animal but not in an ether-anesthetized one. Hunt, Jernberg, and Lawlor (1953) found the same effect using ECS as the convulsing agent. Siegel, McGinnies, and Box (1949) gave ECS following runway training when their animals had received either an injection of neutral saline or a nembutal anesthetic. They found a retention decrement due to ECS when the subjects were injected by the saline but not when under the anesthetic. Each of these studies indicates that RA does not occur when the ECS is given while the subject is under an anesthetic.

Pearlman et al. (1961) in an excellent experiment add a possible complication to the picture by showing that an anesthetic itself can produce RA. They first bar trained animals to a stable rate of responding and then gave them a CER produced by a single shock delivered through the reinforcing apparatus. Either 20 seconds, 5 minutes, 10 minutes, or 20 minutes after the CER the animals received a sodium pentobarbital anesthetic through

chronically implanted catheters. The experimenters found that the anesthetic produced RA as a function of the time between CER and the administration of the anesthetic. In another experiment of the same study they gave pentylenetetrazol, which acts as a convulsant, either 20 seconds, 2 hours, 4 hours, 8 hours, or 4 days after the CER. Here they found that the pentylenetetrazol was effective in attenuating the CER even when given 4 days after the CER. They concluded that the sodium pentobarbital produced RA by interfering with a consolidating engram but that the pentylenetetrazol must have produced forgetting by some other mechanism.

Leukel and Quinton (1964) have added to our knowledge of the effects of ECS in combination with anesthetics. Following the acquisition of a response, they administered carbon dioxide gas after each of several extinction trials. They reasoned that if carbon dioxide disrupted consolidation, the effects of the extinction trials would not be apparent; whatever learning extinction involved would be disrupted. Instead they found that the anesthetic *increased* the rate of extinction, a finding which they took to mean that the anesthetic served as punishment.

REPEATED ECSS AND LEARNING TRIALS

Most of the experiments cited so far in this paper have used repeated learning trials and repeated administrations of ECS. As several experimenters have recently pointed out, these studies have doubtful relevance for consolidation theory. McGaugh and Hudspeth (1964), for example, show that repeated ECSSs can bring about RA that is not due to interference with consolidation. They maintain that only

studies using single ECSSs are relevant to the consolidation process.

When more than one learning trial is given, other difficulties arise. With repeated trials the early engrams have probably already consolidated by the time the later ones are created. For this reason Pearlman et al. (1961) have argued that only situations in which but one learning trial (and one ECS) is given are suitable to test a consolidation process.

If these arguments concerning one-learning trial and one ECS are accepted, it means that there are very few ECS experiments in the literature which support a consolidation notion. The frequently cited Duncan (1949) study, for example, cannot be considered as evidence for consolidation, and all but one of those studies cited by Glickman (1961) must be discarded. The only published ECS experiments, by these criteria, that can even be considered as relevant to a consolidation process are those of Pearlman et al. (1961), Heriot and Coleman (1962), McGaugh and Hudspeth (1964), and Weissman (1963). But even these few studies may not be relevant. Chorover and Schiller (1964), for example, present data which they interpret to show that the consolidation process is completed within 10 seconds after a learning trial. If they are correct, then this is the only ECS study relevant to consolidation.

It seems that the pressure of the ECS data, much of which has been reviewed here, is increasingly restricting the arena in which consolidation may be tested. It is not inconceivable that the possibility of such tests may soon vanish. The kind of evidence that is usually taken as most supportive of consolidation is that which shows RA to be a function of the time between learning and ECS. It is important to note that this type of gradient can be

found in both single- and multiple-trial experiments and appears even when a very long interval separates ECS from the learning trial (Brady, 1952; Williams, 1963).

A CONDITIONED INHIBITION HYPOTHESIS

While the evidence is clear that ECS produces a decrement in the performance of a previously learned response, it is far from clear that the data demand an explanation in terms of the disruption of a storage process. Inasmuch as the problem of amnesia is, by definition, one of forgetting, there is a *prima facie* reason to examine the possibility that ECS produces its effects by processes similar to those seen in other instances of forgetting.

If we take the point of view that unconsciousness is an unconditional response produced by ECS, then the possibility is opened that some weakened version of unconsciousness will become conditioned to whatever external stimuli preceded the onset of the coma. Now coma, or unconsciousness, induced by any method represents a relatively extreme value of inhibition on a dimension of arousal-inhibition. The characteristics of such a dimension are well enough described elsewhere that it is not necessary to elaborate them here. However, high values of inhibition are found in sleep, coma, and deep anesthesia. Lesser values are described in terms of general muscular relaxation, drowsiness, lowered activity levels, and the like.

Many workers in the field of neuropsychology have followed the lead of Pavlov (1957) in regarding deep levels of inhibition as being "protective" in nature, i.e., they serve as a massive protection against further stimulation and excitation when this would be hazardous to the organism. For example,

Jung and Tonnies (1950) have pointed out the central importance of the mechanisms whereby the brain prevents and counteracts massive synchronous convulsive discharges, and we may here regard the coma induced by ECS as an inevitable reaction of protective inhibition to such massive stimulation. Further, we assume that pervasive inhibitory states of this kind are conditionable. At least this would seem to be the case from the observations of Pavlov (1957) upon the rapid development of conditioned sleep in animals exposed to the monotony and restraint of the conditioning apparatus.

In brief, the typical ECS-RA experiment appears to meet the specification for the induction of unconditioned protective inhibition, and we should expect that some weaker magnitude of this state should become conditioned to the surrounding environmental stimuli. By the same token we should expect that the power of a particular environmental stimulus to elicit conditioned inhibition will diminish with its temporal distance from the onset of the unconditioned inhibition. In other words, the usual laws governing the relationship between response strength and CS-UCS interval will hold.

The conditioned inhibition that will be produced by the environmental stimuli accompanying the original ECS will be, as mentioned above, somewhat weaker than complete coma. General muscular relaxation, lowered level of activity, and the like should be found when the animal is in the presence of the CS. Any local stimuli that are particularly related to the delivery of ECS and thus were discriminative stimuli for it should become particularly potent conditioned inhibitors. For example, the present writers have noted informally that, following ECS, the rat will tend to relax and become limp when the ear clips are reattached

for a subsequent trial—in marked contrast to their usual reaction to the first encumbrance of ear clips. By the same token, simple and rather naturalistic observation of rats returned to an apparatus in which they were previously convulsed suggests that the crouching that interferes with avoidance responding is produced by muscular *relaxation*, and is quite different from the tense “freezing” that is seen in the CER pattern.

Conditioned inhibition, like any other conditioned response, should begin to diminish with repeated presentations of the CS, unaccompanied by the UCS. Thus, while the apparatus in which the ECS is delivered should, on next retesting, elicit marked inhibition, this should become progressively less evident with repeated exposure of the apparatus. In this connection Lewis and Adams (1963) describe the effects of such extinction trials on animals returned to an apparatus in which they had previously received ECS: “They were left there for 5 minutes at a time on 5 successive days. Although no formal record was kept of their movements, it was noticeable that they became more and more *active* as this extinction experience continued [p. 516].” They note further that animals receiving this extinction experience more readily learned an active avoidance response later. We shall return to this observation a little later.

That the animals are behaving in a relaxed manner in the face of the CS is further documented by the remarks of Williams (1961) who confirmed the earlier observations of Hunt et al. (1952). Williams pointed out an “interesting paradox of electroconvulsive shock.” She noted that following ECS, the treated animals when in their home cages “huddled in a ball in the rear corners of the cages or crouched on the food containers. They

were hyperirritable when handled and squealed, leaped or ‘froze’ when *E* attempted to lift them [p. 636].” This behavior was in marked contrast to their response when faced with the experimental apparatus in which a CER had been established earlier: “In this experiment, the ECS animals behaved as if the grid were not emotion-arousing [p. 636].” They showed no fear, and exhibited the seeming paradox of an enhancement of emotional responsiveness outside the apparatus but marked attenuation of it in the apparatus.

It is entirely possible that the seeming paradox is resolvable by assuming that the stimuli most associated with ECS now elicit conditioned inhibition, and this state is especially antagonistic to high states of arousal, i.e., the states that prevail in the mediation of emotional behavior. The studies of Adams and Lewis (1962a, 1962b), already cited, are entirely congruent with this inference. When ECS was administered at a location other than that in which the learned response usually occurred, it had less amnesic power than ECS given in the learning apparatus—even though the time interval was the same in both conditions. Adams and Lewis originally conjectured that a partial form of the convulsion was conditioned to stimuli associated with ECS. Thus, RA appeared because the animal was now crouching and huddling instead of performing the original response. One of the few data perplexing to this point of view was the observation that the animals were limp and relaxed upon being returned to the ECS situation; behavior hardly consistent with the notion of a conditioned convulsion. According to the present point of view, the ECS results in inhibition which becomes conditioned to the ECS stimuli. There is, then, a “learned relaxa-

tion," which interferes and competes with the original response and produces RA.

Returning to the literature surveyed in the first part of this paper, we note that the overwhelming bulk of investigations have demonstrated forgetting of avoidance responses, and in particular the forgetting of avoidance responses such as freezing that seem to be mediated by very high states of arousal. It was pointed out that ECS was particularly selective for this type of behavior. This is no longer surprising from an inhibition point of view as we should expect that precisely this kind of response would be most adversely affected by pronounced relaxation—rather analogous to Wolpe's (1958) paradigm of "reciprocal inhibition." The animal cannot freeze if he is relaxed. Responses that are mediated by somewhat lower levels of arousal will be less disrupted and will more readily reappear when the "emotional" response has been antagonized. However, in the absence of more comprehensive investigation of the effects of ECS upon simple approach responses, it is difficult to do more than speculate on this point.

The assumption that ECS produces conditioned inhibition has some power to account for the proactive "amnesic" effects that have been cited earlier. If the new learning situation contains stimulus elements that were present when ECS was delivered (e.g., ear clips attached to the animal, or any general apparatus similarity), then we should expect conditioned inhibition to interfere significantly with the acquisition of a new learned response. Where this similarity was clearly present (Adams & Lewis, 1962) this deduction is indeed confirmed.

In this connection we should note that the effects of ECS are apparently attenuated by prior anesthesia. This is

not surprising. In the first place, the subject, under anesthesia, is already unconscious and at a high level of general inhibition. The subsequent passage of current should be irrelevant and ineffective in conditioning inhibition. It is quite possible to deduce that if the unconsciousness induced by the anesthetic is rapid, and done in the experimental situation, then it should have the same retrograde amnesic effects that ECS alone does. The Pearlman et al. (1961) study seems to meet these specifications and did, as we have seen, produce retrograde amnesic phenomena. The Leukel and Quinton (1964) study is also consistent with this interpretation. They found that an anesthetic speeded up the extinction of a previously learned response, rather than produced RA for the extinction learning. We assume that it was the anesthetic-produced inhibition that increased extinction. The Hunt et al. (1952) study points further to the combined effects of extinction and ECS.

The main tenor of this argument may be summarized as follows:

1. ECS produces massive protective inhibition, a weakened version of which then becomes conditioned to the environmental stimuli present at the time coma was induced.

2. This conditioned inhibition will show itself behaviorally by measurable motor relaxation in the presence of the conditioned stimuli, and this will tend to interfere with the performance of any other response that was previously conditioned to them. Where these latter responses were evoked under high levels of arousal, as for example in "emotional" responding, the attenuation will be especially pronounced.

3. Because this conditioned inhibition is a learned pattern, like any other, it will produce proactive conflict with

new learning just as it produces retroactive interferences with old learning.

4. Where the coma is induced in an environment other than that in which the amnesic effects are to be measured, then these effects will be lessened.

5. The other laws of conditioning will apply to learned inhibition. However, in view of the massive nature of the unconditioned stimulus, effective retention intervals may be expected to exceed the limits usually found in typical laboratory-conditioning procedures.

6. Any procedure such as anesthesia that renders the animal unconscious before passage of current will prevent the establishment of an ECS-produced inhibition or its conditioning to local stimuli. However, if unconsciousness is induced rapidly by anesthetic, even without convulsion, then the same conditioned inhibition should develop to the environment in which the anesthesia was given.

Now none of this, per se, directly touches upon the issue of engram destruction or consolidation. Instead, it attempts to subsume a large number of studies, many of them having somewhat paradoxical relationships with each other, under a single and more general rubric. It seems quite possible that the consolidation notion has not yet been tested at all. If engram formation takes the brief span of time that usually defines immediate memory, then most of the learning-ECS intervals used by investigators are too long to disrupt it. Furthermore, it seems difficult to regard any study in which the subject has had more than one trial as appropriate to the problem, as here again the engram formed by the first trial and other early trials must surely be consolidated by the time the investigator gets to the ECS stage of the investigation.

Ironically, it is the success of investi-

gators in showing "retrograde amnesia" from ECS given many hours, days, or months after the engram was formed that makes it clear that the effects of ECS must be attributed to something other than interference with consolidation. One of the present writers has, with Adams, offered an earlier and tentative hypothesis that the interference effects were produced by conditioning part of the convulsion pattern itself. However, the modification of this view to substitute conditioned inhibition seems better to fit the behavior of the post-ECS animal and also accounts more directly for the antagonism of ECS effects for high-arousal patterns such as the CER.

CONCLUSION

We have reviewed that portion of the ECS literature which has been considered relevant to the problem of neural consolidation. We have seen that very little, if any, of this literature provides convincing support for a consolidation process. This is in no way to deny the existence of such a process, for in many ways it remains appealing. It is simply to say that adequate evidence for consolidation is not available in the ECS literature as has been claimed (Glickman, 1961). We have further presented an interpretation of the effects of ECS in producing RA. This interpretation is based on the concept of conditioned inhibition, and is not meant to exclude the possibility that other mechanisms are also operating. The effects of ECS are undoubtedly very complicated.

REFERENCES

- ADAMS, H. E., & LEWIS, D. J. Electroconvulsive shock, retrograde amnesia, and competing responses. *Journal of Comparative and Physiological Psychology*, 1962, **55**, 299-301. (a)
ADAMS, H. E., & LEWIS, D. J. Retrograde amnesia and competing responses. *Journal*

- of *Comparative and Physiological Psychology*, 1962, 55, 302-305. (b)
- BRADY, J. V. The effects of electroconvulsive shock on a conditioned emotional response: The permanence of the effect. *Journal of Comparative and Physiological Psychology*, 1951, 44, 507-511.
- BRADY, J. V., HUNT, H. F., & GELLER, L. The effect of electroconvulsive shock on a conditioned emotional response as a function of the temporal distribution of the treatments. *Journal of Comparative and Physiological Psychology*, 1954, 47, 454-457.
- BRAUN, H. W., PATTON, R. A., & BARNES, H. W. Effects of electroshock convulsions upon the learning performance of monkeys: I. Object-quality discrimination learning. *Journal of Comparative and Physiological Psychology*, 1952, 45, 231-238.
- CARSON, R. C. The effect of electroconvulsive shock on a learned avoidance response. *Journal of Comparative and Physiological Psychology*, 1957, 50, 125-129.
- CHOROVER, S. L., & SCHILLER, P. H. Short-term retrograde amnesia (RA) in rats. Paper delivered at Eastern Psychological Association, Philadelphia, April 1964.
- DEUTSCH, J. A. Higher nervous function: The physiological bases of memory. *Annual Review of Physiology*, 1962, 24, 259-286.
- DUNCAN, C. P. The retroactive effect of electroshock on learning. *Journal of Comparative Physiology*, 1949, 42, 32-44.
- GELLER, I., SIDMAN, M., & BRADY, J. V. The effect of electroconvulsive shock on a conditioned emotional response: A control for acquisition recency. *Journal of Comparative and Physiological Psychology*, 1955, 48, 130-131.
- GERARD, R. W. Biological roots of psychiatry. *Science*, 1955, 122, 225-230.
- GLICKMAN, S. E. Perseverative neural processes and consolidation of the neural trace. *Psychological Bulletin*, 1961, 58, 218-233.
- HEISTAD, G. T. An effect of electroconvulsive shock on a conditioned avoidance response. *Journal of Comparative and Physiological Psychology*, 1955, 48, 482-487.
- HERIOT, J. T., & COLEMAN, P. D. The effects of electroconvulsive shock on retention of a modified "one-trial" conditioned avoidance. *Journal of Comparative and Physiological Psychology*, 1962, 55, 1082-1084.
- HUNT, H. F., & BRADY, J. V. Some effects of electroconvulsive shock on a conditioned emotional response ("anxiety"). *Journal of Comparative and Physiological Psychology*, 1951, 44, 88-98.
- HUNT, H. F., JERNBERG, P., & BRADY, J. V. The effect of electroconvulsive shock (ECS) on a conditioned emotional response: The effect of post-ECS extinction on the reappearance of the response. *Journal of Comparative and Physiological Psychology*, 1952, 45, 589-599.
- HUNT, H. F., JERNBERG, P., & LAWLOR, W. J. The effect of electroconvulsive shock on a conditioned emotional response: The effect of electroconvulsive shock under ether anesthesia. *Journal of Comparative and Physiological Psychology*, 1953, 46, 64-68.
- HUNT, H. F., JERNBERG, P., & OTIS, L. The effect of carbon disulphide convulsions on a conditioned emotional response. *Journal of Comparative and Physiological Psychology*, 1953, 46, 465-469.
- JOHN, E. R. High nervous functions: Brain functions and learning. *Annual Review of Physiology*, 1961, 23, 451-484.
- JUNG, R., & TONNIES, J. F. Hirnelektrische Untersuchungen über Entleerung und Erhaltung von Kramfentladungen: Die Vorgänge am Reizort und die Bremsfähigkeit des Gehirns. *Archiv für Psychiatrie*, 1950, 185, 701-735.
- KIMBLE, G. A. *Hilgard and Marquis' conditioning and learning*. New York: Appleton-Century-Crofts, 1961.
- KRAINS, S. H. *Mental depressions and their treatment*. New York: Macmillan, 1957.
- LEUKEL, F. A. Comparison of the effects of ECS and anaesthesia on acquisition of the maze habit. *Journal of Comparative and Physiological Psychology*, 1957, 50, 300-306.
- LEUKEL, F., & QUINTON, E. Carbon dioxide effects on acquisition and extinction of avoidance behavior. *Journal of Comparative and Physiological Psychology*, 1964, 57, 267-270.
- LEWIS, D. J., & ADAMS, H. E. Retrograde amnesia from conditioned competing responses. *Science*, 1963, 141, 516-517.
- MCGAUGH, J. L., & MADSEN, M. C. Amnesic and punishing effects of electroconvulsive shock. *Science*, 1964, 144, 182-183.
- MADSEN, M. C., & MCGAUGH, J. L. The effect of ECS on one-trial avoidance learning. *Journal of Comparative and*

- Physiological Psychology*, 1961, 54, 522-523.
- MÜLLER, G. E., & PILZECKER, A. Experimentelle Beiträge zur Lehre vom Gedächtnis. *Zeitschrift für Psychologie*, 1900, Suppl. No. 1.
- PAVLOV, I. P. *Experimental psychology and other essays*. New York: Philosophical Library, 1957.
- PEARLMAN, C. S., JR., SHARPLESS, S. K., & JARVIK, M. E. Retrograde amnesia produced by anesthetic and convulsant agents. *Journal of Comparative and Physiological Psychology*, 1961, 54, 109-112.
- PORTER, P. B. S., & STONE, C. P. Electroconvulsive shock in rats under ether anesthesia. *Journal of Comparative and Physiological Psychology*, 1947, 40, 441-456.
- POSCHER, B. P. H. Proactive and retroactive effects of electroconvulsive shock on approach-avoidance conflicts. *Journal of Comparative and Physiological Psychology*, 1957, 50, 392-396.
- POSTMAN, L. The present status of interference theory. In C. Cofer (Ed.), *Verbal learning and verbal behavior*, New York: McGraw-Hill, 1961.
- SIEGEL, P. S., MCGINNIES, E. M., & BOX, T. C. Runway performance of rats subjected to electroconvulsive shock following nembutal anesthesia. *Journal of Comparative and Physiological Psychology*, 1949, 42, 417-422.
- STONE, C. P., & BAKHTIARI, A. B. Effects of electroconvulsive shock on maze relearning by albino rats. *Journal of Comparative and Physiological Psychology*, 1956, 49, 318-320.
- THOMPSON, R., & DEAN, W. A further study on the retroactive effect of ECS. *Journal of Comparative and Physiological Psychology*, 1955, 48, 488-491.
- THOMPSON, R., & PENNINGTON, D. F. Memory decrement produced by ECS as a function of the distribution of original learning. *Journal of Comparative and Physiological Psychology*, 1957, 50, 401-404.
- UNDERWOOD, B. J. Interference and forgetting. *Psychological Review*, 1957, 64, 49-60.
- UNDERWOOD, B. J., & POSTMAN, L. Extra-experimental sources of interference in forgetting. *Psychological Review*, 1960, 67, 73-95.
- WEISSMAN, A. Effect of electroconvulsive shock intensity and seizure pattern on retrograde amnesia in rats. *Journal of Comparative and Physiological Psychology*, 1963, 56, 806-810.
- WILLIAMS, GERTRUDE J. The effect of electroconvulsive shock on an instrumental conditioned emotional response ("conflict"). *Journal of Comparative and Physiological Psychology*, 1961, 54, 633-637.
- WILLIAMS, GERTRUDE J. The effect of varying the interval between conflict training and electroconvulsive shock on an instrumental conditioned emotional response. *Journal of Comparative and Physiological Psychology*, 1963, 56, 129-131.
- WOLPE, J. *Psychotherapy by reciprocal inhibition*. Stanford: Stanford University Press, 1958.

(Received June 25, 1964)

THEORETICAL NOTES

COMMENT ON THE EXCHANGE OF THEORETICAL NOTES BETWEEN SMITH AND BLACK AND LANG

ROBERT B. MALMO

Allan Memorial Institute, McGill University

In an exchange of notes Black and Lang present previously unpublished muscle-potential data indicating the relative effectiveness of curare drugs in (at least) greatly reducing skeletal-motor accompaniments of cardiac conditioning. This weakens the theoretical position of Smith who holds that ANS conditioning is an artifact. In my note I comment on the muscle-potential data of Black and Lang and refer to additional data from cardiac conditioning experiments with brain stimulation that are also interpreted as weakening Smith's theoretical position. In conclusion, a combined experimental attack on the problem is suggested.

A recent exchange of theoretical notes (Black & Lang, 1964; Smith, 1964a, 1964b) was concerned

with the question of whether or not curare drugs . . . forestall responses in skeletal muscles [Smith, 1964a].

One of the main theoretical issues was one earlier expounded by Smith (1954). This is the general problem of whether or not autonomic nervous-system (ANS) conditioning can occur as a primary phenomenon or whether it is invariably an artifact of a truly primary reaction in the skeletal motor system. In the first note Smith (1964b) expressed reservations concerning certain conclusions drawn in two recent articles (Black, Carlson, & Solomon, 1962; Solomon & Turner, 1962) in which curare drugs had been employed with the intention of effectively eliminating the influence of skeletal motor response mediation of conditioning. The paper by Black, Carlson, and Solomon (1962) dealt specifically with ANS conditioning. Smith's reservations had mainly to do with the actual effectiveness of the curare drugs in reducing skeletal motor reactions to such a low level that they could not reasonably be expected to be the primary mediators of the ANS conditioned responses.

In Black and Lang's (1964) reply, muscle-potential data were presented

showing that the curare drugs had indeed greatly reduced the level of background muscle tension. In fact, at the amplifier gains used the muscle potentials failed to appear in the tracings. In Smith's (1964a) reply he stated that the muscle-potential data were far more convincing than the unaided observations of overt responses reported earlier, and he indicated that he would be willing to abandon his earlier theoretical position if further studies of this kind continued to provide data contrary to this position.

My note will deal with two points: (a) a technical point concerning the relative adequacy of the muscle-potential data; and (b) another line of attack on Smith's theoretical position, different from that of experimentation with curare drugs.

Black and Lang's (1964) muscle-potential data represent a real technical advance, but as they realized¹ their data would be much more informative were they to employ an amplifier with a good frequency response up to 200 cycles per second and with considerably more gain, such that good recording could be obtained at the level of 1 microvolt per millimeter.

My second point, as previously mentioned, is concerned with a different line

¹ A. H. Black, personal communication, 1964.

of attack on Smith's (1954) position: bypass the sensory pathways and stimulate close to the "motor area" for the ANS reaction, thus avoiding peripheral stimulation (e.g., electric shock) that unavoidably produces strong concomitant skeletal motor reactions (in the absence of curare drugs). A recent experiment (Malmo, 1963, p. 26 ff.) has demonstrated the feasibility of this brain-stimulation approach. With lateral septal stimulation as unconditioned stimulus in a classical conditioning experiment, the unconditioned heart-rate slowing reaction was conditioned to a tone in a group of 20 rats. Respiration was ruled out as an artifact (Malmo, 1963, p. 31 ff.; 1964). From available evidence it was also concluded that the heart-rate slowing reactions were probably not artifacts of skeletal motor reactions to brain stimulation. Although such reactions were not studied as systematically as respiratory reactions (Malmo, 1963, 1964), it was clear that when movements did follow brain stimulation (some animals showed little or no reaction), they varied widely in kind and in intensity from subject to subject, always, however, being associated with heart-rate slowing (for methods employed see Malmo, 1963, p. 26 ff.; and for published data bearing on the question of possible skeletal motor artifacts see Malmo, 1963, p. 18 ff. and p. 31).

In short, the evidence from this experiment appeared strongly to support the conclusion that the heart-rate reactions, conditioned and unconditioned, were primary and not in any way artifacts of skeletal movement. Again, the major positive finding was that in contrast to the widely varying patterns of respiratory and skeletal motor reactions from subject to subject the heart-rate reactions were consistently in the direction of slowing. The evidence strongly supported the conclusion that septal stimulation elicited an autonomic reaction that was a primary reaction, and furthermore that this reaction was conditioned.

The artifact notion of Smith (1954) seems extremely difficult to defend in the presence of these data, especially when they are considered together with those of Black and Lang (1964). If at higher gain muscle potentials still fail to appear and conditioning occurs when curare drugs are injected, using the procedure of Black and Lang, the results of course would be quite conclusive. On the other hand, in case there is some slight residual muscle activity revealed at high gain, the peripheral shock stimulus could be replaced by central stimulation in order to produce heart-rate slowing with little or no activation of the skeletal motor system; and under these more favorable conditions, injection of a curare drug should completely eliminate all muscle potentials.

REFERENCES

- BLACK, A. H., CARLSON, N. J., & SOLOMON, R. L. Exploratory studies of the conditioning of autonomic responses in curarized dogs. *Psychological Monographs*, 1962, 76(29, Whole No. 548).
- BLACK, A. H., & LANG, W. M. Cardiac conditioning and skeletal responding in curarized dogs. *Psychological Review*, 1964, 71, 80-85.
- MALMO, R. B. On central and autonomic nervous system mechanisms in conditioning, learning, and performance. *Canadian Journal of Psychology*, 1963, 17, 1-36.
- MALMO, R. B. Heart rate and respiratory changes in self-stimulation of dorsal and lateral septal areas in the rat. Paper read at Eastern Psychological Association, Philadelphia, April 1964.
- SMITH, K. Conditioning as an artifact. *Psychological Review*, 1954, 61, 217-225.
- SMITH, K. Comment on the paper by Black and Lang. *Psychological Review*, 1964, 71, 86. (a)
- SMITH, K. Curare drugs and total paralysis. *Psychological Review*, 1964, 71, 77-79. (b)
- SOLOMON, R. L., & TURNER, LUCILLE H. Discriminative classical conditioning in dogs paralyzed by curare can later control discriminative avoidance responses in the normal state. *Psychological Review*, 1962, 69, 202-219.

(Received August 7, 1964)

IN DEFENSE OF A NEW APPROACH TO OLD PHENOMENA

NORMAN J. SLAMECKA

University of Vermont

The soundness of Dallett's (1965) correction procedures as applied to data from the method of association was recognized, but his conclusions about the existence of remote associations were disputed on the ground that they lacked statistical confirmation. Bugelski's (1965) arguments and experimental data were evaluated and judged as inadequate. It was concluded that the doctrine of remote associations is still unsupported and that alternative explanations are stronger.

Evidently the doctrine of remote associations (RAs) can still summon some supporters to its tattered banner. Dallett (1965) and Bugelski (1965) have risen to the challenge implicit in my paper (Slamecka, 1964) and have offered some interesting arguments and data in defense of the traditional doctrine. Their remarks, together with this reply, should inspire wider scrutiny of an aspect of associationism which has for too long been taken for granted.

Dallett confined his remarks solely to the treatment of data obtained by the method of association. He described a rational ad hoc procedure which corrects for opportunity, and also for stimulus and response biases in such data (the latter being only an approximate correction), and eventuates in an expected gradient which is to be compared to the actually obtained gradient of presumed RAs. He applied this procedure to the data of my Experiment IV, which involved no serial learning, and also to some serial learning based data of his own and presented the resulting gradients.

Presumably, if those elusive RAs are actually operating, a comparison of the two gradients would show a systematic difference between them. At low degrees of remoteness the obtained values should lie well above the expected and should then slope steadily and progressively downward until they lie well below the expected values at the highest degrees of remoteness, thus generating a monotonic gradient whose slope clearly exceeds that of the expected gradient. This is a per-

fectly acceptable rationale which constitutes a real advance in methodology.

However, when Dallett (1965) then presented the gradients from his own data and claimed that "inspection of Figure 2 reveals that the obtained RA gradient is in general steeper than the expected one [p. 167]," I could no longer agree. It reveals no such thing. The difficulty with his conclusion is that it was really unsupported. No objective test was provided by which to arrive at a valid comparison of the gradients. Mere visual inspection, accompanied by some subjective judgmental process, is a totally unacceptable method for evaluating those particular curves. The courses of the obtained points throughout his Figure 2 show erratic reversals in every case rather than the steady, monotonic declines to be hoped for. Further, in the graphing of the background gradients the "-1" degree (whose values extend upward so extremely far) should not even have been included, because that degree represents adjacent associations and not remotes. The differences among the remaining frequencies appear relatively unimpressive; and since the standard errors of such data may be relatively large, it is a very real possibility that the curves lie within the limits of random variation.

Dallett's chief contribution to the issue is his sophisticated analysis of the sources of bias that surround the method of association; but, as far as the doctrine of remote associations is concerned, the emperor is still wearing no clothes.

Bugelski's (1965) critical remarks were so extensive that I have chosen to reply only to those that seemed most pertinent. These points will be taken up in the order in which they came in his paper.

1. Bugelski's assertion that I assume the serial position curve "requires no explanation," is false, as is his assertion that I was begging the question. The "question" at hand concerned the validity of the evidence surrounding the existence of RAs, and it was far from being begged. My statement (Slamecka, 1964) was,

Whatever the ultimately accepted explanation of the serial position curve will be (and papers are still being published about that problem), it is unlikely that remote associations will play any part in it [p. 74].

In fact, recent evidence suggests that the serial position curve may be based upon the learning of relative item positions (Ebenholtz, 1963a; Jensen, 1962) rather than upon some complex bundle of presumed RAs.

2. Through the use of some rather unconvincing arguments, Bugelski attempted to shore up the old assumption that serial learning is a matter of forming direct chain associations. However, his presentation is somewhat out of keeping with the newer experimental facts and serves mainly to illustrate the powerful hold of the Ebbinghaus tradition. The findings of such students as Ebenholtz (1963b), Underwood (1963), and Young (1962) make it quite clear that the stimulus for an item in a serial list is not the preceding item or even a cluster of preceding items.

My suggestion that serial learning proceeds by fixing the relative positions of the items was a recognition of the general trend of those findings. I quite agree that if one were asked to name the seventeenth letter of the alphabet without counting, he would probably miss it. I also feel that one would not answer with a, b, c, d, e, or v, w, x, y, z, but would come reasonably *close* to the seventeenth letter. It is undeniable that in serial learning one does, in fact, get to know quite a bit about the ordinal positions of

the items (Schultz, 1955), but further reflection indicates that this is not the central aspect of the process. Thanks to Bugelski's clarity in putting the question, the answer can also be made clearer. I believe that it is the *relative* positions of the items that are being learned, and only incidentally their ordinal positions. Thus, the appropriate type of question to ask our subject would be: "What letter appears later in the alphabet, j or m, e or l, p or t?" and the like. I predict that one could answer such relational questions with near-perfect accuracy. No amount of juggling with hypothetical RAs could account for the feat.

3. Apparently Bugelski is willing to concede that the derived list technique is practically useless for providing real proof of the validity of RAs, and I agree wholeheartedly. Nevertheless, the logic of that technique *demands* that a first-order list (having the strongest "remote" associations) be learned faster than a scrambled control list (with necessarily weaker associations on the average). My first three experiments showed that such an outcome occurs only if the list is a regularly patterned one and (a) is either derived from a familiar sequence such as the alphabet, or (b) the subject is aware of the principle of patterned derivation.

Bugelski also complained that the use of an alphabet list tends to "confuse the issue." Since the reason for using an alphabet sequence in my Experiment I was not well taken, one need only consult Experiments II and III which made the same basic points with nonsense syllables! These experiments provided strong support for a perception-of-patterning rationale and gave no comfort to the remote associations doctrine.

4. Bugelski had two objections to my experimental demonstration that the overall results of the method of association are artifacts of response biases and can be produced without any serial learning whatever. First, he felt that a pre-familiarization frequency of 25 for the first item and of 1 for the fourth item was not a good analogue of serial learning, since it implies that 25 trials

are needed to master a six-syllable list. Second, he was alarmed over the finding of such a high proportion of backward associations, in spite of the fact that previous studies have always reported a substantial number of them. Even Dallett (1959) reported more backward-remote than forward-remote associations at each of his three retention intervals. Would Bugelski also consider those data "fantastic"?

In order to indicate that neither of the two objections is crucial to the issue, Table 1 is presented. It depicts some hypothetical data from the method of association, illustrating a perfect response bias from the serial position effect during original learning. There is a ratio of 2 to 1 between the first-item and fourth-item responses. Since in my Experiment IV a ratio of more than 4 to 1 was obtained when the respective practice frequencies were 25 and 1, it follows (assuming a logarithmic relation between trials and free-response probabilities) that the six-item list of Figure 1 might be learned in about seven trials. In this table there are more forward than backward associations (90-72), more adjacent forward than adjacent backward (27-22), and more remote forward than remote backward across degrees 1 through 4 (58-50). Having satisfied Bugelski's requirements for acceptability, these data

will still show the usual pattern of RAs as a function of degree of remoteness and thus indicate that the objections were not really important as a criticism of the response-bias rationale.

5. Bugelski felt there was an inconsistency in my position that (a) with the method of association RAs are artifacts of response overlearning (due to differential practice frequencies), and (b) with the method of anticipatory and perseverative errors they are artifacts of item underlearning (insufficient learning of item positions). This observation is superficial since it overlooks the realistic aspects of the two situations.

With the method of association the subject is free to emit any list response, and therefore the determining influence of prior frequency is allowed full rein. With the method of anticipatory and perseverative errors the subject is engaged in serial learning, and his responses are necessarily more constrained. He is trying to anticipate the next item in the list, so he refrains from uttering those that have already passed and those whose identity and positions are known well enough to be recognized as incorrect at that point. The influence of serial position is evident in the progressively narrowing range of the list from which intrusion errors come. Such constraints do not exist in the former, free-responding situation. These explanations are ad hoc, and properly so, since they are aimed at two quite different and specific testing conditions.

6. Now for consideration of the one-trial experiment which supposedly rebutted my entire position on the method of anticipatory and perseverative errors with one fell swoop. Bugelski presented a serial list to classroom groups for one exposure only and then presented the third word and asked for the one that followed it. Half of the answers were correct, and the remainder were distributed in progressively decreasing frequencies as their distance from the correct word increased. Bugelski assumed that I would predict a random distribution of intrusions, so he viewed the result as contrary to my position. The second part

TABLE 1
HYPOTHETICAL DATA OF THE METHOD OF
ASSOCIATION ILLUSTRATING A PERFECT
RESPONSE BIAS

Stimulus items	Response item (in serial order)					
	1	2	3	4	5	6
(X)	6	5	4	3	4	5
1	—	5	4	3	4	5
2	6	—	4	3	4	5
3	6	5	—	3	4	5
4	6	5	4	—	4	5
5	6	5	4	3	—	5
6	6	5	4	3	4	—

Note.—The number in each cell is the frequency with which that response was made to the given stimulus during the association test.

of the study was an unaided serial recall test of the entire list. Bugelski assumed that I would predict a negative correlation between the number of items correctly learned as to position, and the extent to which they occurred as errors in the first part of the study. The obtained correlation was positive, so it was again viewed as contradicting my position.

These data do not constitute a rebuttal of my position. Bugelski's rationale for the study was evidently based upon a gross oversimplification of my account of the process of serial learning. It must be remembered that two stages were postulated: a response-learning phase and a position-learning phase. I also said (Slamecka, 1964),

the subject . . . occasionally makes a guess as to the next item, uttering any one he happens to *know* [*italics added*] except those he has just seen . . . and those . . . whose identity and position have been fixed [p. 72].

One cannot offer a response which he has not yet learned. A glance at the last column of Bugelski's Table 2 shows that the response-learning phase favored the early items, and that there was a systematic decline in item availability as the list progressed. Items 7 through 10 went unreported by more than half of the subjects—obviously they were not all equally available as intrusions! The use of common words as items does not eliminate the response-learning phase from consideration at all.

Under such conditions I would certainly not predict a random distribution of errors, but would instead expect the very distribution that was obtained, namely, one that mirrored the pattern of response availabilities. The first two items, whose identities and positions were both well learned, were rarely given as intrusions, again according to expectations. Even if all of the items were equally learned, I would predict a random-error distribution *only* if there had been no position learning accomplished at all. The moment that position learning showed itself, there would be a restriction of the

range of intrusions according to the serial position function.

This same differential response learning renders Bugelski's correlation data confounded with respect to its interpretation and makes his subsequent conclusions untenable. The only condition under which a negative correlation between intrusion frequency and degree of position learning should be found would be when all the items were equally available. The correlation would simply reflect the progress of the position-learning phase, and its magnitude should increase as the degree of position learning increased.

An appropriate method of testing my position within the framework of Bugelski's own study would be to provide the subject with a scrambled listing of the items for reference during the recall tests. This would guarantee equal item availability and would allow the results to reflect position learning unconfounded by differential response learning. With such a procedure a curvilinear distribution of errors should result, as well as the above-mentioned negative correlation.

After having evaluated the remarks of these critics, I am more than ever convinced that there is less than meets the eye in the doctrine of RAs, and that the tide is surely ebbing for the house of Ebbinghaus.

REFERENCES

- BUGELSKI, B. R. In defense of remote associations. *Psychological Review*, 1965, 72, 169-174.
- DALLETT, K. M. Retention of remote associations. *Journal of Experimental Psychology*, 1959, 58, 252-255.
- DALLETT, K. M. In defense of remote associations. *Psychological Review*, 1965, 72, 164-168.
- EBENHOLTZ, S. M. Position mediated transfer between serial learning and a spatial discrimination task. *Journal of Experimental Psychology*, 1963, 65, 603-608.
- (a)
- EBENHOLTZ, S. M. Serial learning: Position learning and sequential associations. *Journal of Experimental Psychology*, 1963, 66, 353-362. (b)
- JENSEN, A. R. Temporal and spatial effects of serial position. *American Journal of Psychology*, 1962, 75, 390-400.

- SCHULTZ, R. W. Generalization of serial position in rote serial learning. *Journal of Experimental Psychology*, 1955, **49**, 267-272.
- SLAMECKA, N. J. An inquiry into the doctrine of remote associations. *Psychological Review*, 1964, **71**, 61-76.
- UNDERWOOD, B. J. Stimulus selection in verbal learning. In C. N. Cofer & B. S. Musgrave (Eds.), *Verbal behavior and learning*. New York: McGraw-Hill, 1963.
- YOUNG, R. K. Tests of three hypotheses about the effective stimulus in serial learning. *Journal of Experimental Psychology*, 1962, **63**, 307-313.

(Received September 1, 1964)

PSYCHOLOGICAL REVIEW

SPATIAL VARIABLES, OBSERVING RESPONSES, AND DISCRIMINATION LEARNING SETS¹

FRED STOLLNITZ

Brown University

An observing response is any response that results in exposure to a discriminative stimulus. 2 assumptions are made: that the probability of occurrence of an observing response varies (a) directly with cue area, and (b) inversely with spatial separation between cue and response. With these assumptions, observing-response theory explains the effects of these 2 spatial variables on discrimination learning by monkeys, chimpanzees, and children. Prolonged training on a single discrimination problem can overcome the difficulty produced by these variables, but, surprisingly, the difficulty persists through extensive learning-set training. It is concluded that changing problems every few trials in learning-set training results in extinction of any observing response that might be reinforced within problems.

When a rhesus monkey is trained for a few trials on each of many discrimination problems, it forms a discrimination learning set (Harlow, 1949, 1959). That is, its performance gradually changes from practically no progress on each problem to rapid solution of each new problem. How rapid? One-trial discrimination of

three-dimensional objects is well known. In fact, learning-set formation has been described as "converting a problem which is initially difficult for a subject into a problem which is so simple as to be immediately solvable [Harlow, 1949, p. 56]." This view was supported by the results of Harlow and Warren (1952), who found that rhesus monkeys (*Macaca mulatta*) could form a learning set when trained with 3-inch squares cut from the pages of magazines. The initial great difficulty of these discriminations was gradually overcome, and eventually many problems were solved in single trials.

So it was quite reasonable for Blazek and Harlow (1955) to expect that performance differences between initially easy and initially difficult problems would gradually disappear during learning-set formation. They trained

¹Based in part on a PhD dissertation (Stollnitz, 1963a) written while I was a National Science Foundation Graduate Fellow. My experiments and the preparation of this paper were supported by United States Public Health Service Research Grants M-2818 and MH-07136 from the National Institute of Mental Health. Allan M. Schrier, who directed the dissertation, gave patient advice, helpful criticism, and kind encouragement. Comments by Kenneth V. Anderson, Donald S. Blough, Judith L. Crooks, Charles S. Harris, Judith R. Harris, Donald R. Meyer, V. J. Polidora, Elizabeth M. Stearns, and L. Benjamin Wyckoff have improved this paper.

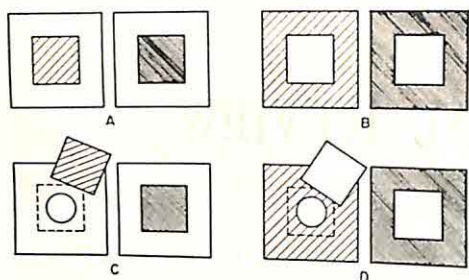


FIG. 1. Four spatial variants of a color-discrimination problem. (A, whole-card center cue; B, whole-card border cue; C, part-card center cue; D, part-card border cue. The colored areas—shaded—cover 25% of the base card in A and C, and 75% in B and D. C and D are shown with a food well uncovered after a response.)

rhesus monkeys to discriminate between differently colored squares centered on 3-inch-square white cards (Figure 1, Part A). Problem difficulty was varied by using colored squares of different sizes: The colored squares covered 100%, 50%, 25%, or 12.5% of the 3-inch-square bases, with both stimulus cards of a given problem having the same size colored squares. The results surprised even the experimenters. As expected, performance improved progressively on all of the size variants as the monkeys formed a learning set. But the monkeys never learned to do as well with the small colored squares as with the large ones. Even after 768 six-trial problems, performance levels remained directly related to the size of the colored squares. The differences among performance levels increased at the start of the experiment but were constant over the last 576 problems. From these results, Blazek and Harlow (1955) concluded that more training could not be expected to reduce the differences in performance.

In a similar experiment, Schrier and Harlow (1956) trained Java monkeys (*Macaca irus*) on 432 10-trial problems with colored squares cover-

ing 100%, 38%, or 25% of 3-inch-square white cards. Performances on 25% and 38% problems were identical but persistently worse (by about 8% correct responses) than performance on 100% problems.

The difficulty of problems like those shown in Figure 1, Part A, was discovered 2 decades ago (Harlow, 1945), but the variables that make them difficult have been identified only recently. This paper will show that the persistent performance differences found by Blazek and Harlow (1955) and Schrier and Harlow (1956), as well as other effects of the same variables, can be explained by an elaboration of Wyckoff's (1952) theory about the role of observing responses in discrimination learning. It should be noted that the present paper does not attempt to explain learning-set formation, but only to show why certain results have been obtained in a number of learning-set experiments. Conceptual developments followed experimental discoveries, of course, but the reverse path will be followed here to avoid some dead ends. The historic byways have been traced by Meyer, Treichler, and Meyer (1965).

OBSERVING-RESPONSE THEORY

Wyckoff (1952) introduced the observing-response concept as a convenient way to deal with attending to discriminative stimuli in discrimination learning experiments. An *observing response* is defined here as *any response that results in exposure to a discriminative stimulus.*² "Ex-

² This definition is more widely applicable than "any response which results in exposure to the pair of discriminative stimuli involved [Wyckoff, 1952, p. 431]." With Wyckoff's definition, an observing response in a successive-discrimination experiment would have to last for two or more stimulus presentations. Moreover, his definition is inappropriate for at least some simulta-

posture" implies that the discriminative stimulus is exteroceptive; thus, observing responses can be distinguished from other chained responses that produce only proprioceptive stimuli (Blough, in press). But an observing response is still a link in a chain leading to the response on which reinforcement is based (the *effective response*). So acquisition and extinction of observing responses are determined by the same variables that determine acquisition and extinction of other chained responses.

In the rest of this paper, "response" means "effective response" unless "observing response" is specified. "Cue" means "discriminative stimulus," as its everyday use suggests. Following Wyckoff (1952), p_o represents the probability of occurrence of an observing response.

If the subject is not exposed to a cue, differential responding is obviously impossible. Therefore, every discrimination task requires an observing response for correct performance, even if the observing response is simply standing with the eyes open. Paradoxically, the situations in which an observing response is most likely to occur are those in which its necessity is most likely to be ignored. Even Wyckoff (1952, p. 431) fell into this trap when he wrote of "situations where no observing response is required of S , that is to say, . . . situations in which S is certain to be exposed to the discriminative stimuli on each trial or prior to each effective response ($p_o =$

1)." When $p_o = 1$, an observing response is just as necessary for correct performance as when p_o is very low. The difference is that when $p_o = 1$ at the beginning of training, no observing response must be *acquired* by the subject, but when p_o is initially very low, an observing response must be acquired before the subject can perform correctly. The present paper is mainly concerned with situations in which p_o is very low at the beginning of training.

The following mechanism for acquisition of an observing response was proposed by Wyckoff (1952): "Exposure to discriminative stimuli will have a reinforcing effect on the observing response to the extent that S has learned to respond differently to the two discriminative stimuli [p. 435]." Later, Zeigler and Wyckoff (1961) found that either learned or unlearned differential responses to discriminative stimuli can reinforce an observing response. Thus, a discriminative stimulus does not have to be paired with reinforcement in order to be a reinforcer (cf. Estes, 1958; Guthrie, 1959; Premack, 1959; Wyckoff, 1954). The "cue strength" (Wyckoff, 1959), "signal value" (Berlyne, 1960, p. 87), or probability of response to the stimulus (Premack, 1959) determines the reinforcing effect on the observing response.

In the visual-discrimination experiments to be discussed, appropriate fixation and accommodation are assumed to be necessary for exposure to a cue.³ So a chain of observing

neous-discrimination experiments, as well. For instance, retarded children often observe only one of a pair of objects on each trial after they have learned the discrimination (Zeaman & Shepp, 1964). Similarly, highly trained monkeys (*Macaca nemestrina*), solving simultaneous-discrimination problems in one trial, observe only the object that they push aside on that trial (Lockhart, Parks, & Davenport, 1963).

³They may also be sufficient for situations in which the same visual dimensions are always correlated with reward and non-reward, as is true of the experiments to be discussed. To handle other situations, Zeaman and House (1963) have postulated that particular dimensions are covertly observed or attended to. Similarly, attending to non-visual cues sometimes requires covert observing responses.

responses may include locomotion, head movements, eye movements, and accommodation movements of the ciliary muscles. For practical purposes, accommodation is a covert response. Eye movements may be measured more directly (e.g., White & Plum, 1964), but this is not convenient when a monkey is running around a cage. So experiments may require arbitrary, easily measured observing responses to be added to the chain (Kelleher, 1958; Wyckoff, 1952; Zeigler & Wyckoff, 1961). Because each link in the chain is reinforced by the following link, all observing responses in the chain are affected similarly by variables that affect the probability of visually fixating a cue. Therefore, knowledge about arbitrary, experimenter-defined observing responses can be used to analyze situations in which the observing response is not measured.

SPATIAL DETERMINANTS OF OBSERVING-RESPONSE PROBABILITY

Separation between Cue and Response

Monkeys look where they put their fingers. Color-discrimination learning by macaque monkeys (*Macaca mulatta* and *M. irus*) is impaired when the cues are placed only 0.50 to 0.75 inch away from where the monkeys touch the display in making their choices (Schuck, 1960; Schuck, Polidora, McConnell, & Meyer, 1961). To account for this influence of spatial separation, Schuck (1960) proposed the existence of "a gradient of visual sampling, with a maximum in the region of the monkey's hand position in manipulation [p. 255]."

The sampling gradient is interpreted here as a spatial distribution of p_o for cues at various distances from the response locus. The quantitative properties of this distribution are not yet

known. Discovering them would probably be made easier if separations between cue and response were stated in terms of visual angle and distance in depth.

A few properties of the distribution are known, however. For example, there is evidence that the distribution is not identical for all directions of cue-response separation. Separations in depth of about 1 inch do not impair performance (Peterson & Rumbaugh, 1963), presumably because the appropriate change in accommodation is much more probable than an appropriate change in fixation for a lateral separation of the same size. But even when cue and response lie in a plane practically perpendicular to the line of sight, the effects of equal cue-response separations in different directions may differ (Horel, Schuck, & Meyer, 1961). Moreover, when there is more than one response locus (as in simultaneous discrimination), the distributions around the several loci may differ. But, for a given response locus and direction of separation, it is assumed that as the separation increases, p_o decreases. The farther the cue from the response locus, the less likely the monkey is to look at it. If the cue is at the response locus, then p_o is maximal.

Cue Area

In the problems used by Blazek and Harlow (1955) and Schrier and Harlow (1956), decreased area of the centered color cues was confounded with increased separation between cue and response (assuming response to borders of cards—see next section). On the other hand, increased area was confounded with increased separation in the experiment by Schuck et al. (1961). Because a 4.75-square-inch increase in area failed to eliminate the effects of introducing a 0.50- to 0.75-

inch cue-response separation, Schuck et al. (1961) concluded that area has "negligible" effects. But this conclusion should be justified only if performance were unrelated to area *with constant cue-response separation*. Experiments holding separation constant (Leary & Lynn, 1961; Warren, 1953; Zimmermann, 1961, 1962; Zimmermann & Torrey, 1965), described below, have shown that cue area does affect color discrimination by rhesus monkeys. Therefore, for a given cue-response separation and at least for color cues, it is assumed here that p_o increases as the cue area increases. Again, probably the visual angle subtended is most important. Obviously, if the cue is so small that a monkey can just barely detect it, then p_o is just barely greater than zero: Only a very precise fixation would be an observing response. If the area is very large, any change of direction of gaze is likely to "hit" the cue.

LEARNING-SET STUDIES

Because the learning-set procedure is so popular in experiments on discrimination learning by monkeys, most of the available evidence relevant to the observing-response theory comes from experiments in which each of many problems was presented for a constant small number of trials (usually 6 to 10). In each experiment, a single group of monkeys was trained with the several spatial variants (treatments-by-subjects design), unless two groups are mentioned.

In the Wisconsin General Test Apparatus, monkeys uncover food wells by pushing aside cards, plaques, or other objects. When the entire surfaces of two such objects differ in color, for example, the monkey must touch the colored surface in pushing the object. Cue and response are contiguous, p_o is high, and the color dis-

crimination is learned rapidly. Now consider a pair of cards or plaques with only their centers colored differently (Figure 1, Part A). If the total area of the cards remains the same (typically 9 square inches), the cue area is smaller. In addition, assuming that monkeys usually touch only the white borders in pushing these cards, cue and response are spatially separated. Decreased area and increased separation have the same effect: p_o is lower than before, and discrimination learning is retarded. This difference has been found repeatedly (Blazek & Harlow, 1955; Harlow, 1945; Schrier & Harlow, 1956, 1957; Warren, 1953). Most of these studies have varied the size of the center cues and obtained performance levels that vary directly with size, as would be predicted both from the changes in area and from the changes in separation. The two spatial variables are confounded in these experiments, and additional evidence is needed to show that *both* cue area and cue-response separation in fact influence learning-set performance.

Cue Area

Still assuming that monkeys touch the borders of cards that they push from food wells, spatial separation between cue and response becomes zero if the centers of the cards are white and the color cues are on the borders (Figure 1, Part B). Nevertheless, Warren (1953) found that decreasing the cue area (decreasing p_o) impaired the performance of rhesus monkeys on such border-cue problems.

The influence of cue area has been confirmed by an experiment in which the place where the monkeys touched the display was known rather than assumed. Leary and Lynn (1961) modified a test tray to hold stimulus cards over the food wells in such a way

that rhesus monkeys could move the cards only by pushing on the near edges. Colored 1-inch squares and 1 × 3-inch rectangles were placed at equal distances from the near edges of the cards. With cue-response separations thus held constant, performance was better with the larger cue areas than with the smaller ones, supporting the present analysis.

Separation between Cue and Response

Macaque monkeys perform better on border-cue problems (Figure 1, Part B) than on center-cue problems (Figure 1, Part A) when the cue area is the same (Riopelle, Wunderlich, & Francisco, 1958; Warren, 1953). This shows that small cue area is not the only cause of difficulty in center-cue problems. As mentioned above, cue-response separation is presumably zero in border-cue problems, so p_0 should be greater for border cues than for center cues of equal area.

Two experiments in which rhesus monkeys had to push on the near edges of stimulus cards similarly indicate the importance of cue-response separation (Leary & Lynn, 1961; Schuck, 1960). In both studies, colored regions at the near edges were discriminated more accurately than colored regions of the same area at the far edges. Cue-response separation was clearly responsible: Cue position had no effect when response locus was uncontrolled (Lynn & Leary, 1962) or when monkeys touched both the near and far edges of the cards (Schuck, 1960).

In the studies described so far, the monkeys displaced the whole card that covered the food well. But Schrier and Harlow (1957) cleverly presented center-cue problems to one group of rhesus monkeys as "whole-card" problems (Figure 1, Part A) and to another group as "part-card" problems

(Figure 1, Part C). Before the monkey responds, these variants look identical. But in part-card problems, the base cards are fastened to the test tray, and the monkey pushes aside only the center square to uncover the food well. Performance was better on part-card center-cue problems than on whole-card center-cue problems because the monkeys touched the colored portions in part-card problems. But when the part-card group was switched to whole-card problems, they continued to perform well, presumably because they continued to touch the colored portions of the cards. Similar transfer of touching the cues facilitated form discrimination by monkeys in a learning-set experiment by Polidora and Fletcher (1964).

Schuck (1960) found that performance of rhesus monkeys was not reliably different on center-cue and border-cue variants given as part-card problems (Figure 1, Parts C and D). In pushing the center squares, the monkeys had their hands near both border and center regions, so p_0 was high for both border and center cues.

To summarize 8 years of experimentation with the four kinds of color-discrimination problem shown in Figure 1: The whole-card center-cue problem (Figure 1, Part A) has consistently been more difficult than the others for macaque monkeys. It is the only one of the four in which p_0 would be substantially lowered if, as is proposed here, cue-response separation were an important determinant of p_0 .

If response to the border of a display results in higher p_0 for border cues than for center cues, then response to the center of a display should result in higher p_0 for center cues than for border cues. This prediction, made by Schuck et al. (1961), was handsomely verified. When macaque monkeys responded by touching buttons

centered on the stimulus screens of an automatic apparatus, they performed better on center-cue problems than on border-cue problems. The difference in performance on the two spatial variants increased during the experiment and then was roughly constant (about 10% correct responses) over the last 448 of 832 six-trial problems.

The persistence of performance differences with differences in cue-response separation was puzzling. It seemed strange that an appropriate observing response could not be acquired to overcome the effects of a cue-response separation of only 0.50 to 0.75 inch, as in the border-cue problems of Schuck et al. (1961). The effects of much larger separations have been overcome in other experiments, to which we now turn.

SINGLE-PROBLEM STUDIES

So far a static view of p_o has been adequate. To account for the effects of spatial variables in learning-set studies, p_o could be assumed to remain throughout each experiment at the initial value determined for each spatial variant by cue area and cue-response separation. But, in contrast to the results of extensive learning-set training, extensive training on a single problem can overcome the difficulty imposed by relatively large spatial separations between cue and response. In other words, observing responses that are not acquired during learning-set formation are acquired during training on single problems. Thus, in this section, spatial variables can be assumed to determine p_o only at the start of training, and changes in p_o during training must be considered.

The importance of cue area in single-problem performance has been demonstrated in experiments with infant rhesus monkeys (Zimmermann, 1961, 1962; Zimmermann & Torrey, 1965).

In these experiments, the monkeys responded to nursing-bottle nipples that were always centered in the cue surface, so cue-response separation was constant. Color discriminations were learned much more rapidly with large cues (high p_o throughout) than with small cues (low initial p_o). Also, after a color discrimination had been learned with a large cue area, reducing the cue area (reducing p_o) impaired performance.

The rate of learning a single discrimination is inversely related to the distance between cue and response loci, as observing-response theory would predict. This is true for chimpanzees (Gellermann, 1933a; Jenkins, 1943) and children (Jeffrey & Cohen, 1964; Lipsitt, 1961; Murphy & Miller, 1959) as well as for rhesus monkeys (Carr & Brown, 1960; McClearn & Harlow, 1954; Meyer, Polidora, & McConnell, 1961; Miller & Murphy, 1964; Murphy & Miller, 1958). In Murphy and Miller's (1958) study, monkeys learned a black-white discrimination within 400 trials when they pushed aside identical objects just below the cue objects, but failed to learn within 400 trials when the cue-response separation was 7 inches. This result was confirmed by Stollnitz and Schrier (1962). But, after various amounts of additional training, some of Stollnitz and Schrier's rhesus monkeys suddenly learned the discrimination with the 7-inch separation. This sudden learning indicates that, for at least some monkeys, the initial value of p_o was greater than zero with a 7-inch cue-response separation. Wyckoff (1952) described a positive-feedback relation between discrimination and observing-response learning that explains the sudden learning, as follows: Discrimination learning was retarded for quite a long time because p_o was so low at first. But learning,

once started, was accelerated positively because every increase in differential responding resulted in greater reinforcement of later observing responses. A similar explanation can be given for learning by some of the children trained by Murphy and Miller (1959) with a 6-inch cue-response separation. In several of these studies, some subjects failed to learn at all; presumably these subjects' initial p_0 was extremely low (perhaps zero).

Rhesus monkeys trained by Stollnitz and Schrier (1962) with small cue-response separations were able to transfer their discrimination to successively greater separations (see also Schrier, Stollnitz, & Green, 1963). Increasing the separation 4 inches or more in one jump seriously impaired performance, but even with an 18-inch separation some of these monkeys (and some previously trained only with the 7-inch separation) made 95% to 98% correct responses. These results extend those of McClearn and Harlow (1954), whose monkeys reached over 90% correct responses with 0-, 1-, 2-, and 4-inch separations presented repeatedly. In both experiments, the discrimination learned with one separation could be transferred to larger separations. One of Murphy and Miller's (1958) monkeys also showed some transfer from the 0- to the 7-inch separation when it continued to touch the positive cue before making the effective response. According to the present theory, each time the separation is increased, p_0 decreases. But even though an observing response is at first less likely to occur, every time one occurs it is reinforced, because the cues still control the differential responding that was learned with smaller separations. So p_0 increases faster when a discrimination is transferred to a larger separation than it would if the discrimination

were being newly learned with that separation.

As would be expected, relatively small separations between cue and response impair learning less severely than do large separations. In fact, when the same problem is given both with no separation and with small separations to the same subjects, the effects of spatial variables may be very small, because the initial depression of p_0 is counteracted so rapidly by the transfer effect just described. This apparently occurred in an experiment by Meyer (1958), who trained rhesus monkeys on a single color-discrimination problem with cues centered on 3-inch-square bases. With five cue areas ranging from 9.0 square inches (100%) to 0.2 square inch (2.2%) presented repeatedly, transfer from the 100% variant to the four center-cue variants made the performance differences small. Meyer concluded that "reducing the area does not substantially increase the difficulty of the problem [1958, p. 183]." But analysis of variance⁴ (sessions \times variants \times subjects) showed that the effect of the spatial variables (area and separation combined) was reliable ($p < .025$).

The combined effect of cue area and cue-response separation was isolated from the transfer effect in an experiment⁵ by Stollnitz (1963a). Rhesus monkeys were trained on a single whole-card color-discrimination problem, using 3-inch-square plaques with their entire surfaces colored (100% variant) or with colored squares centered on white backgrounds (12.5% variant). To eliminate transfer between the variants, they were pre-

⁴ Donald R. Meyer kindly provided the raw data.

⁵ Done while I was a Predoctoral Research Fellow of the National Institute of Mental Health, United States Public Health Service.

sented to two different groups of monkeys. Practically errorless performance was reached after about 100 trials with the 100% variant and after about 200 trials with the 12.5% variant. Meanwhile, a third group was trained with the two variants alternating every 5 trials. After a total of about 100 trials, the alternating group also was making practically no errors on the 100% variant, even though this training included only about 50 trials with the 100% variant. (The other 50 were with the 12.5% variant.) Similarly, after about 100 trials with each variant, the alternating group performed as accurately on the 12.5% variant as did the 12.5 group after about 200 trials with that variant alone. So, although the combined effect of area and separation was still reliable, there was perfect transfer between the two spatial variants in terms of amount of training required for errorless performance.

Even with perfect transfer and constant cue area, small separations between cue and response reliably retard learning of a single color discrimination. Stollnitz (1963a, 1963b) trained rhesus monkeys to discriminate the colors of narrow strips (0.2 inch \times 3.0 inches) on 3-inch-square white plaques that could be pushed only straight back. Cue-response separation was varied at random from trial to trial by placing the colored strip at one of 15 distances from the near edge of each plaque. The total number of errors was linearly related to the distance of the strip from the near edge ($p < .01$), although errorless performance was eventually achieved with all 15 distances. Apparently the discrimination was learned first with the small separations and transferred rapidly to the larger separations as observing responses were acquired for

cues at greater distances from the response locus.

The results of all these experiments would lead one to expect that a 0.5-inch cue-response separation would reliably retard color-discrimination learning by rhesus monkeys, although the effect might be small. So Polidora and Fletcher (1964), like Meyer (1958), may have overlooked such an effect when they reported that the median numbers of days to criterion with and without a 0.5-inch separation were within 1 day of each other. They gave 64 trials a day until 12 consecutive responses were correct within 1 day's session, so the median numbers of trials to criterion may, in fact, have differed by as much as 116 ($64 + 64 - 12$).

With or without transfer, then, spatial variables that reduce p_o increase the difficulty of single discrimination problems. But long training on a single problem can overcome the difficulty, even leading to errorless performance on a whole-card center-cue problem (Stollnitz, 1963a).

PERSISTENCE OF EFFECTS IN LEARNING-SET FORMATION

Why is it that prolonged learning-set training does not eventually reduce or eliminate the differences in difficulty caused by spatial variables? Blazek and Harlow (1955) offered no answer to this question. Neither did Schuck et al. (1961), although they saw that a satisfactory explanation must allow for the fact that the differences can be eliminated by prolonged training on a single problem. They suggested that

control of observing responses in these situations is by some visuomotor process that is highly resistant to modification through generalized, interproblem learning of the kind involved in learning sets. But changes in these processes can and do accompany the

learning involved in the mastery of a single problem [p. 545].

This statement merely pushes the question into the realm of "some visuo-motor process." Why does this "process" resist change during learning-set formation and not during single-problem learning? To put the question more directly, why does p_o remain at or near its initial value throughout a learning-set experiment?

The easiest answer would be that p_o never changes from its initial value *within* any problem, and so necessarily remains constant (for a given spatial variant) throughout the experiment. This answer is most obviously true if p_o is zero. If the cues are never observed, observing can never be reinforced, so p_o never changes. But even if p_o is greater than zero, at least two observing responses must be made before p_o could increase. The second observing response could be reinforced only if the subject had learned to respond differentially when it made the first observing response. Even assuming one-trial discrimination learning, then, any very low initial value of p_o may result in a negligible chance for p_o to increase within a short problem. In fact, with very short problems and very low p_o , there would be little chance even for discrimination learning. Such an experiment was done by Murphy and Miller (1955), who trained rhesus monkeys on three-trial object-quality discrimination problems. With a 6-inch cue-response separation, one group showed no sign of learning during 672 problems. The other group, responding by pushing aside the stimulus objects, formed a learning set but fell to chance performance when switched to the 6-inch separation after 480 problems. Presumably the 6-inch separation did not reduce p_o quite to zero, since rhesus monkeys can learn

a *single* discrimination with a 7-inch separation (Stollnitz & Schrier, 1962).

But p_o does not have to be very low to remain constant within every problem. Consider a spatial variant for which p_o is extremely high. For example, in the typical "100%" color-discrimination problem, the subject must touch the colored surface if it responds at all. Because primates look where they put their fingers, p_o is close to 1.0 at the start of discrimination training. The same is true of object-quality discrimination when subjects push aside the stimulus objects. Of course, the fact that p_o is close to 1.0 does not imply that discrimination performance starts out nearly perfect, but with an initially high p_o the discrimination is learned relatively fast. As a learning set is formed, new problems are solved more and more rapidly, but p_o remains essentially unchanged. It cannot rise above 1.0 no matter how much the observing response is reinforced, and it cannot fall below its initial value (close to 1.0) as long as the subject touches the cues. In this sense, the spatial conditions set a lower limit on p_o .

What if the initial value of p_o is neither extremely high nor extremely low? For example, what if p_o is .50 at the start of discrimination training? Note first that a subject could perform 75% correct responses on a two-choice discrimination with a p_o of .50 by responding correctly every time the cues were observed and half the time (chance level) when they were not. So the simplest assumption still is that p_o remains constant within problems while discrimination performance improves. If training on a single problem were continued for a long time, p_o would eventually start to increase after the discrimination had been learned enough for the observing response to be reinforced. But in a learning-set

experiment, learning is cut off after 10 trials or even fewer. Warren's (1953) within-problem learning curves show that performance differences among spatial variants are still present on Trial 10. Enough observing responses have been made by then to permit some discrimination learning, but the problem is changed before p_o can increase. Discrimination learning must then begin once again for the new pair of cues. So, from the simple assumption that p_o remains constant within one problem, it follows that p_o remains constant throughout the experiment.

That simple assumption may not always be the most plausible, however. The higher the initial value of p_o , the faster each problem would be solved if training were continued, and the fewer trials would be needed on a given problem before the observing response would be reinforced and p_o would start to increase. So it could be argued that p_o might increase a little within a short problem, especially after a learning set has been formed and considerable discrimination learning is occurring within each problem. But now what happens when the problem is changed? Recall that an observing response will be reinforced only to the extent that the subject responds differentially to the cues. In a learning-set experiment, the cues of each new problem do not at first control differential responding. Exposure to the new cues will fail to reinforce the observing response. Even if the value of p_o has increased slightly during one problem, it will fall again as the observing response extinguishes when the next problem is started. So p_o will not increase progressively from problem to problem, and the effects of spatial variables will persist throughout the experiment.

Fortunately, the operation of such a mechanism has been demonstrated. In

an experiment by Zeigler and Wyckoff (1961), pigeons had to step on a pedal to expose themselves to cues. This observing response was acquired during each of six discrimination problems, but was found to extinguish after each new problem was introduced. The repeated acquisition and extinction confirmed Zeigler and Wyckoff's prediction from observing-response theory and directly supports the present analysis.

Should one conclude that p_o cannot increase progressively during learning-set formation? High initial values of p_o , long problems, and fast-learning subjects might foster such an increase. But, to explain the persistence of the effects of spatial variables in the experiments reviewed here, it seems plausible to claim that any observing response that might have been reinforced within problems extinguished again when the problem was changed. Thus p_o was held close to its initial level by this repeated extinction. After all, the first trial of a new problem is always an extinction trial for the observing response, so extinction always has a head start.

DISCUSSION

Extinction of the observing response when a new problem is introduced has been a hidden help to those who have trained monkeys on a number of spatial variants of color-discrimination problems using the learning-set procedure and the treatments-by-subjects design. Because such extinction prevents transfer among variants when the cues are changed, this economical design has revealed the same spatial effects as would be shown if a different group of subjects was trained on each variant.

The theory presented here to explain the effects of spatial variables is formally similar to Restle's (1958)

theory of within-problem learning in learning-set experiments. In Restle's theory, irrelevant stimuli ("type-*c* cues") lose their control of behavior ("adapt") in the presence of ("with respect to") the relevant cues of a particular problem ("type-*b* cues"). As a result, the irrelevant stimuli gain control again (are "released from adaptation") whenever the relevant cues are changed. In observing-response theory, reinforcement of an observing response by exposure to particular cues results in extinction of any acquired observing response whenever these cues are changed. Some of Restle's concepts may be identified with the empirical variables that have been discussed in this paper. For example, when cue and response are spatially separated, there must always be some irrelevant stimuli between them. If these stimuli are identified with Restle's type-*c* cues, then adaptation of these cues with respect to the relevant cues becomes functionally equivalent to acquisition of an observing response for the relevant cues. This approach suggests the possibility of applying Restle's logic to an observing-response theory of learning-set formation. Restle accounts for learning-set formation by postulating that both type-*c* and type-*b* cues adapt with respect to type-*a* cues (cues that are valid throughout the experiment). This mechanism might correspond to acquisition of an (covert?) observing response for the cues of reward and nonreward on the previous trial. Such an observing response, of course, would occur in addition to the observing response for the relevant cues, which has been the main concern of this paper.

It is interesting to notice why Restle thought that his model of learning-set formation could not deal with observing-response learning. Finding that

his model fit the data of Harlow and Warren (1952) poorly, Restle suggested that the monkeys in that experiment were learning observing responses that directed their fixations toward the centers and away from the borders of the stimulus cards. But that was the experiment (mentioned in the introduction of this paper) which used 3-inch squares cut from magazines, which presumably differed as much in their borders as in their centers. The discriminations were difficult, but not because the monkeys had to fixate the centers of the squares.

Spatial variables are important in a wide variety of situations. For example, facilitation of correct performance by transfer of a cue-touching response has been mentioned earlier in connection with experiments by Schrier and Harlow (1957), Murphy and Miller (1958), and Polidora and Fletcher (1964). Perhaps through transfer of preexperimental learning, children and chimpanzees trained by Gellermann (1933a, 1933b) with cues separated from the response loci "spontaneously" traced the contours of form cues with their fingers before responding. Instructions to trace contours with the fingers similarly facilitated form discrimination by children in an experiment by Ruzskaia (1958), and tracing words helps children to discriminate them in learning to read (Fernald, 1943). All of these seem to be similar cases of primates looking where they put their fingers, rather than (as has often been claimed) of kinesthetic cues supplementing or modifying visual perception. Other mammals, lacking fingers, look where they put their noses. Thus, touching the cues seems to be quite generally important for efficient discrimination performance (Ehrenfreund, 1948; Gardner & Nissen, 1948; Munn, 1931)—even rats look before they leap.

Atkinson (1961) wrote that more experimental data were needed before the concepts of observing-response theories could be related to empirical variables, such as the properties of stimuli. At about the same time, Schuck et al. (1961) recognized that the effects of spatial properties of stimuli should be explained in terms of observing-response theory. This paper is an attempt to satisfy both requests.

REFERENCES

- ATKINSON, R. C. The observing response in discrimination learning. *Journal of Experimental Psychology*, 1961, **62**, 253-262.
- BERLYNE, D. E. *Conflict, arousal, and curiosity*. New York: McGraw-Hill, 1960.
- BLAZEK, NANCY C., & HARLOW, H. F. Persistence of performance differences on discriminations of varying difficulty. *Journal of Comparative and Physiological Psychology*, 1955, **48**, 86-89.
- BLOUGH, D. S. The study of animal sensory processes by operant methods. In W. K. Honig (Ed.), *Operant behavior: Areas of research and application*. New York: Appleton-Century-Crofts, in press.
- CARR, R. M., & BROWN, W. L. Association of relevant stimuli along a spatial gradient. *Journal of Genetic Psychology*, 1960, **97**, 131-137.
- EHRNFREUND, D. An experimental test of the continuity theory of discrimination with pattern vision. *Journal of Comparative and Physiological Psychology*, 1948, **41**, 408-422.
- ESTES, W. K. Stimulus-response theory of drive. In M. R. Jones (Ed.), *Nebraska symposium on motivation: 1958*. Lincoln: Univer. Nebraska Press, 1958. Pp. 35-69.
- FERNALD, GRACE M. *Remedial techniques in basic school subjects*. New York: McGraw-Hill, 1943.
- GARDNER, L. PEARL, & NISSEN, H. W. Simple discrimination behavior of young chimpanzees: Comparisons with human aments and domestic animals. *Journal of Genetic Psychology*, 1948, **72**, 145-164.
- GELLERMANN, L. W. Form discrimination in chimpanzees and two-year-old children: I. Form (triangularity) *per se*. *Journal of Genetic Psychology*, 1933, **42**, 3-27. (a)
- GELLERMANN, L. W. Form discrimination in chimpanzees and two-year-old children: II. Form versus background. *Journal of Genetic Psychology*, 1933, **42**, 28-50. (b)
- GUTHRIE, E. R. Association by contiguity. In S. Koch (Ed.), *Psychology: A study of a science*. Vol. 2. New York: McGraw-Hill, 1959. Pp. 158-195.
- HARLOW, H. F. Studies in discrimination learning by monkeys: III. Factors influencing the facility of solution of discrimination problems by rhesus monkeys. *Journal of General Psychology*, 1945, **32**, 213-227.
- HARLOW, H. F. The formation of learning sets. *Psychological Review*, 1949, **56**, 51-65.
- HARLOW, H. F. Learning set and error factor theory. In S. Koch (Ed.), *Psychology: A study of a science*. Vol. 2. New York: McGraw-Hill, 1959. Pp. 492-537.
- HARLOW, H. F., & WARREN, J. M. Formation and transfer of discrimination learning sets. *Journal of Comparative and Physiological Psychology*, 1952, **45**, 482-489.
- HOREL, J. A., SCHUCK, J. R., & MEYER, D. R. Effects of spatial stimulus arrangements upon discrimination learning by monkeys. *Journal of Comparative and Physiological Psychology*, 1961, **54**, 546-547.
- JEFFREY, W. E., & COHEN, LESLIE B. Effect of spatial separation of stimulus, response, and reinforcement on selective learning in children. *Journal of Experimental Psychology*, 1964, **67**, 577-580.
- JENKINS, W. O. A spatial factor in chimpanzee learning. *Journal of Comparative Psychology*, 1943, **35**, 81-84.
- KELLEHER, R. T. Stimulus-producing responses and attention in the chimpanzee. *Journal of the Experimental Analysis of Behavior*, 1958, **1**, 87-102.
- LEARY, R. W., & LYNN, ELIZABETH. Discrimination by rhesus monkeys of patterns with asymmetrically placed cues. *Psychological Reports*, 1961, **9**, 361-368.
- LIPSITT, L. P. Simultaneous and successive discrimination learning in children. *Child Development*, 1961, **32**, 337-347.
- LOCKHART, J. M., PARKS, T. E., & DAVENPORT, J. W. Information acquired in one trial by learning-set experienced monkeys. *Journal of Comparative and Physiological Psychology*, 1963, **56**, 1035-1037.
- LYNN, ELIZABETH, & LEARY, R. W. Reinforcement procedure and cue location in pattern discriminations of monkeys. *Psychological Reports*, 1962, **11**, 83-90.
- MCCLEARN, G. E., & HARLOW, H. F. The effect of spatial contiguity on discrimina-

- tion learning by rhesus monkeys. *Journal of Comparative and Physiological Psychology*, 1954, 47, 391-394.
- MEYER, D. R. Some psychological determinants of sparing and loss following damage to the brain. In H. F. Harlow & C. N. Woolsey (Eds.), *Biological and biochemical bases of behavior*. Madison: Univer. Wisconsin Press, 1958. Pp. 173-192.
- MEYER, D. R., POLIDORA, V. J., & MCCONNELL, D. G. Effects of spatial S-R contiguity and response delay upon discriminative performances by monkeys. *Journal of Comparative and Physiological Psychology*, 1961, 54, 175-177.
- MEYER, D. R., TREICHLER, F. R., & MEYER, PATRICIA M. Discrete-trial training techniques and stimulus variables. In A. M. Schrier, H. F. Harlow, & F. Stollnitz (Eds.), *Behavior of nonhuman primates*. Vol. 1. New York: Academic Press, 1965. Pp. 1-49.
- MILLER, R. E., & MURPHY, J. V. Influence of the spatial relationships between the cue, reward, and response in discrimination learning. *Journal of Experimental Psychology*, 1964, 67, 120-123.
- MUNN, N. L. An apparatus for testing visual discrimination in animals. *Journal of Genetic Psychology*, 1931, 39, 342-358.
- MURPHY, J. V., & MILLER, R. E. The effect of spatial contiguity of cue and reward in the object-quality learning of rhesus monkeys. *Journal of Comparative and Physiological Psychology*, 1955, 48, 221-224.
- MURPHY, J. V., & MILLER, R. E. Effect of the spatial relationship between cue, reward, and response in simple discrimination learning. *Journal of Experimental Psychology*, 1958, 56, 26-31.
- MURPHY, J. V., & MILLER, R. E. Spatial contiguity of cue, reward, and response in discrimination learning by children. *Journal of Experimental Psychology*, 1959, 58, 485-489.
- PETERSON, MARJORIE E., & RUMBAUGH, D. M. Role of object-contact cues in learning-set formation in squirrel monkeys. *Perceptual and Motor Skills*, 1963, 16, 3-9.
- POLIDORA, V. J., & FLETCHER, H. J. An analysis of the importance of S-R spatial contiguity for proficient primate discrimination performance. *Journal of Comparative and Physiological Psychology*, 1964, 57, 224-230.
- PREMACK, D. Toward empirical behavior laws: I. Positive reinforcement. *Psychological Review*, 1959, 66, 219-233.
- RESTLE, F. Toward a quantitative description of learning set data. *Psychological Review*, 1958, 65, 77-91.
- RIOPELLE, A. J., WUNDERLICH, R. A., & FRANCISCO, E. W. Discrimination of concentric-ring patterns by monkeys. *Journal of Comparative and Physiological Psychology*, 1958, 51, 622-626.
- RUZSKAIA, A. G. Orienting-investigatory activity in the formation of elementary generalizations in the child. In L. G. Voronin et al. (Eds.), *Orientirovochny refleks i orientirovochno-issledovatel'skaia deiatel'nost'*. [The orienting reflex and exploratory behavior.] Moscow: Academy of Pedagogical Sciences, 1958. Cited by D. E. Berlyne, *Conflict, arousal, and curiosity*. New York: McGraw-Hill, 1960. P. 218.
- SCHRIER, A. M., & HARLOW, H. F. Effect of amount of incentive on discrimination learning by monkeys. *Journal of Comparative and Physiological Psychology*, 1956, 49, 117-122.
- SCHRIER, A. M., & HARLOW, H. F. Direct manipulation of the relevant cue and difficulty of discrimination. *Journal of Comparative and Physiological Psychology*, 1957, 50, 576-580.
- SCHRIER, A. M., STOLLNITZ, F., & GREEN, K. F. Titration of spatial S-R separation in discrimination by monkeys (*Macaca mulatta*). *Journal of Comparative and Physiological Psychology*, 1963, 56, 848-851.
- SCHUCK, J. R. Pattern discrimination and visual sampling by the monkey. *Journal of Comparative and Physiological Psychology*, 1960, 53, 251-255.
- SCHUCK, J. R., POLIDORA, V. J., MCCONNELL, D. G., & MEYER, D. R. Response location as a factor in primate pattern discrimination. *Journal of Comparative and Physiological Psychology*, 1961, 54, 543-545.
- STOLLNITZ, F. *Performance of monkeys (Macaca mulatta) on spatial variants of color-discrimination problems*. (Doctoral dissertation, Brown University) Ann Arbor, Mich.: University Microfilms, 1963, No. 64-2011. (a)
- STOLLNITZ, F. Spatial gradient of performance of monkeys (*Macaca mulatta*) on a single color-pattern discrimination problem. *American Psychologist*, 1963, 18, 473. (Abstract) (b)
- STOLLNITZ, F., & SCHRIER, A. M. Discrimination learning by monkeys with spatial separation of cue and response. *Journal of Comparative and Physiological Psychology*, 1962, 55, 876-881.

- WARREN, J. M. The influence of area and arrangement on visual pattern discrimination by monkeys. *Journal of Comparative and Physiological Psychology*, 1953, **46**, 231-236.
- WHITE, S. H., & PLUM, G. E. Eye movement photography during children's discrimination learning. *Journal of Experimental Child Psychology*, 1964, **1**, 327-338.
- WYCKOFF, L. B., JR. The role of observing responses in discrimination learning: Part I. *Psychological Review*, 1952, **59**, 431-442.
- WYCKOFF, L. B., JR. A mathematical model and an electronic model for learning. *Psychological Review*, 1954, **61**, 89-97.
- WYCKOFF, L. B. Toward a quantitative theory of secondary reinforcement. *Psychological Review*, 1959, **66**, 68-78.
- ZEAMAN, D., & HOUSE, BETTY J. The role of attention in retardate discrimination learning. In N. R. Ellis (Ed.), *Handbook of mental deficiency*. New York: McGraw-Hill, 1963. Pp. 159-223.
- ZEAMAN, D., & SHEPP, B. E. Some peripheral correlates of attention in retardate discrimination learning. Paper read at Eastern Psychological Association, Philadelphia, April 1964.
- ZEIGLER, H. P., & WYCKOFF, L. B. Observing responses and discrimination learning. *Quarterly Journal of Experimental Psychology*, 1961, **13**, 129-140.
- ZIMMERMANN, R. R. Analysis of discrimination learning capacities in the infant rhesus monkey. *Journal of Comparative and Physiological Psychology*, 1961, **54**, 1-10.
- ZIMMERMANN, R. R. Discrimination of colored surfaces by neonatal and adult rhesus monkeys. Paper read at Eastern Psychological Association, Atlantic City, April 1962.
- ZIMMERMANN, R. R., & TORREY, C. C. Ontogeny of learning. In A. M. Schrier, H. F. Harlow, & F. Stollnitz (Eds.), *Behavior of nonhuman primates*. Vol. 2. New York: Academic Press, 1965. Pp. 405-447.

(Received April 30, 1964)

NEW DEVELOPMENTS IN FACET DESIGN AND ANALYSIS¹

URIEL G. FOA

Israel Institute of Applied Social Research, Jerusalem

In multivariate research design the systematic definition of the set of variables in terms of more basic sets, the facets, leads to the prediction of the empirical interrelationship among the variables. 2 principles are suggested for predicting the results from the facet structure of the variables: the principle of contiguity and the semantic principal components. The principle of contiguity simply states that variables more similar in their facet structure will also be more related empirically. This principle does not provide for a differentiation among facets in determining the relationship. A nonparametric approach to the differentiation problem is provided by the more general concept of semantic principal component. The principle of contiguity turns out to be the special case of the first principal component. The application of these concepts to a number of studies in different behavioral areas suggests that they have predictive power. It is further shown that facet elements can be classified into specific and nonspecific to the set of variables and that variables containing specific elements tend to be related to the set of variables more than variables containing nonspecific elements. Systematic design alone does not guarantee correct prediction of empirical results. In fact, for a given area of behavior several alternative formalizations appear possible, and they will usually lead to different hypotheses. While the choice of a given facet design rather than another may depend on the intuition of the investigator, it appears also to be related to the psychology of concept information and to the influence of language on this process.

Behavioral studies tend to become wider in scope and to include a larger number of variables. The increasing complexity of data requires, more than ever, systematic design and a rigorous statement of hypotheses. Facet metatheory is an approach to the treatment of this problem of multivariate design and analysis. It is in-

¹ The work on general methods of research design and analysis described here has been supported in part by the Office of Aerospace Research, United States Air Force, through its European Office, under Contract No. AF 61(052)-121. Its application to substantive studies has been facilitated by Grant M-2669 from the National Institute of Mental Health, National Institutes of Health, U. S. Public Health Service.

The preparation of this paper was made possible by a grant from the Lucius N. Littauer Foundation, New York.

tended to help and guide the intuition of the investigator rather than to substitute for it. A scientific concept is constructed or invented by abstracting certain features from the data of experience and disregarding certain other aspects. In deciding which aspects should be considered and which should not, the investigator has to rely initially on his intuition. The same universe of experience can be conceptualized in many different ways, depending on the choice of the investigator; and some events may belong to the same or to different concepts according to the features one chooses to consider.

Concepts become accepted by others, or become "scientific," insofar as they lend themselves to neat empirical

relations. One of the difficulties besetting behavioral science at present is precisely that different investigators make different choices, thus precluding the possibility of comparing the findings they obtain. Investigators are apparently spurred to try new concepts by the fact that concepts previously used rarely yield neat and orderly data. A necessary criterion (albeit not a sufficient one) for a good theory is that it should lead systematically to the correct prediction of empirical results. Facet design and analysis suggest an approach to such a testing.

The concepts presented in this paper have been developed in specific areas of study. The presentation will however stress the formal rather than the substantive aspects in the hope that this will facilitate the application of this approach to other areas. But, again, there is no intention to suggest that a research problem can be approached from a purely formal viewpoint. The methodology of design and analysis and substantive theory are closely entwined, and no research problem seems likely to be solved by considering one aspect only (Brunswik, 1951; Hammond, 1951).

FACET DESIGN

The idea that variables included in a research design should be defined in terms of their component elements is not new. Its use in behavioral research was stimulated by the desire to apply multivariate factorial design to this field (Fisher, 1949). In the early 1930s Brunswik and Reiter employed a factorial design in studying the impression value of schematized faces (Brunswik, 1956, pp. 100-108). They used 189 facial schemes, each scheme being defined by different combinations of elements of the forehead, eyes, nose, and mouth. For example, one of the faces is defined by high forehead, wide

eye separation, medium-tip and short-saddle nose, high mouth. Brunswik (1956) discusses at length the problems related to systematic and representative design and gives many other examples of its use, drawn mainly from the field of perception.

A similar approach is proposed by Stephenson (1953). One of the examples given by him (see pp. 69-70) is a redefinition of Jung's psychological types in terms of Attitude (introversion, extraversion), Mechanism (conscious, unconscious), and Function (thinking, feeling, sensation, intuition). This leads to $2 \times 2 \times 4 = 16$ types. One type, for example, is "unconscious extravert feeling."

Guttman (1954b) suggested that this manner of defining variables can be formalized by using the notation of the Cartesian product and introduced the term *facet* to indicate a component set of the product. By doing this, Guttman opened the way for further developments.

To illustrate this notation let us use again Stephenson's example of Jungian types. The three facets are: A, Attitudes, with two elements: a_1 , introversion, and a_2 , extraversion; B, Mechanism, with two elements: b_1 , conscious, and b_2 , unconscious; and C, Function, with four elements: c_1 , thinking, c_2 , feeling, c_3 , sensation, and c_4 , intuition. The Cartesian product ABC indicates the set of 16 types; the notation $a_1b_1c_1$ indicates one element of the product, namely, conscious introvert thinking; other elements are $a_1b_1c_2$, $a_1b_2c_1$, and so on.

Some of the variables thus defined are more similar among themselves than some others. Variable $a_1b_1c_1$, for example, is more similar to $a_1b_1c_2$ than to $a_2b_2c_2$. The first two variables have the elements of Facets A and B in common, while the first and last variables have no common elements.

This particular way of defining psychological types or faces, as in Brunswik's study, is only one of the many possible alternatives. Different definitions can be thought of, using different facets; and these different definitions will show different similarity patterns. The process of constructing a concept consists in retaining certain facets and disregarding certain other facets from the data of experience. Facet design formalizes this intuitive approach by spelling out the facets which are retained, but it does not tell, *a priori*, which facets should be spelled out in the definition. The choice of facets is a substantive rather than a methodological problem. Facet analysis permits, however, a test whether a particular facet design produces similarity patterns which are confirmed by empirical results. Thus the conceptual structure of the variables may suggest hypotheses about their statistical relationship.

PRINCIPLE OF CONTIGUITY

A possible relationship between the conceptual and empirical structure of the variables is proposed by the principle of contiguity (Foa, 1958). This principle states that variables which are more similar in their facet structure will also be more related empirically. Using this principle in our example we shall predict that the relationship between $a_1b_1c_1$ and $a_1b_1c_2$ will be higher than the relationship between $a_1b_1c_1$ and $a_2b_2c_2$.

The use of the notion of contiguity or proximity or similarity for predicting the relationship among variables had been limited in the past to the rather special case in which the variables could be located along some physical dimensions, such as loudness of sound, brilliancy of light, color, position in the physical space, and the like. More often, in the absence of a clearly defined variable space, this notion was

resorted to in order to explain the found relationship rather than for predicting it. An illustration of this explanatory use is given by Guttman (1954a, especially p. 339). In the explanatory argument the existence of a similarity pattern is inferred from the empirical relationship. On the other hand, in the principle of contiguity the empirical relationship is predicted from the similarity pattern of facet elements. In the facet context proximity acquires a precise meaning. It is now proximity in terms of facet elements. Thus it becomes possible to use it for predicting the relationship rather than for explaining it.

The first indication that the principle of contiguity may have some predictive power was obtained in a study of the interpersonal relations between 490 factory workers and their 51 foremen. In this study the relationship between 32 interpersonal variables was analyzed. It was found that, on the average, variables with more facet elements in common correlate higher than variables with fewer common elements. The average correlations were as follows: variables with elements of two facets in common, .23; variables with the element of one facet in common, .10; variables with no common elements, -.01 (Foa, 1958).

In fact this study goes farther than proposing the principle of contiguity. It also suggests the more specific hypothesis that variables having more facet elements in common will be more related than variables having fewer facet elements in common. The principle of contiguity is necessary for this hypothesis to be true, but not sufficient. To support it, it is also necessary that the various facets should have approximately the same weight in determining the relationship. If Facet A contributes to the relationship much more than Facets B and C, it may hap-

pen that variables differing only in the elements of Facet A will be less related than variables differing in the elements of both B and C.

The specific hypothesis of counting common facet elements was derived from the more general principle of contiguity (Foa, 1958). The distinction between these two notions was not maintained, however, in the analysis of the facet structure of a simplex (Guttman, 1959).

THE FACET STRUCTURE OF A SIMPLEX

In this work Guttman (1959) re-analyzes a study of intergroup beliefs and action of 580 white Brazilian respondents toward the Negro. Three facets were used:

Facet A. The *level* of intergroup behavior with two elements: a_1 , belief, and a_2 , overt action.

Facet B. The *type* of intergroup behavior with two elements: b_1 , comparative, and b_2 , interactive.

Facet C. The *referent* to which the intergroup behavior is ascribed: c_1 , the subject's group, and c_2 , the subject himself.

The Cartesian product of these three facets defines variables such as the following ones:

1. $a_1b_1c_1$ Stereotype: *belief* (a_1) of a subject that *his own group* (c_1) (ex-cels—does not excel) in comparison (b_1) to Negroes.

2. $a_1b_2c_1$ Norm: *belief* (a_1) of a subject that *his own group* (c_1) (ought—ought not) to *interact* (b_2) with Negroes.

3. $a_1b_2c_2$ Hypothetical interaction: *belief* (a_1) of a subject that *he himself* (c_2) (will—will not) *interact* (b_2) with Negroes.

4. $a_2b_2c_2$ Personal interaction: *Overt action* (a_2) of the *subject himself* (c_2) (to—not to) *interact* (b_2) with Negroes.

These four variables are ordered by similarity or contiguity: $a_1b_2c_1$ is the most similar to $a_1b_1c_1$, differing only in the element of Facet B. Next comes $a_1b_2c_2$ which differs from $a_1b_1c_1$ in the element of Facets B and C. Likewise $a_1b_2c_2$ is the best neighbor of $a_1b_2c_1$ and of $a_2b_2c_2$, and so on. Indeed the four variables so ordered form a perfect Guttman scale.

The principle of contiguity now suggests that Variable 1 will empirically relate most highly to Variable 2, less to Variable 3, and least to Variable 4. Variable 2, in turn, will be more related to Variables 1 and 3 than to Variable 4, and so on.

Scores, ordered from favorable to unfavorable to the Negroes, were assigned to the 580 white respondents in the four variables, on the basis of their answers to questions pertaining to each particular variable. These scores were then intercorrelated. The intercorrelation coefficients confirmed the prediction, with only one slight deviation.²

The correlation pattern obtained is a *simplex*, a statistical structure which is characterized by the possibility of ordering the variables along a line. Those nearer on the line correlate among themselves higher than those more distant (Guttman, 1954a). It may be of some interest to note that contiguity alone is not sufficient for predicting the simplex. It has to be supplemented by the more severe criterion of counting facet elements. To give an example let us compare the rela-

² For details on this point see the technical appendix to this paper which has been deposited with the American Documentation Institute. Order Document No. 8249 from ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington, D. C. 20540. Remit in advance \$1.75 for microfilm or \$2.50 for photocopies, and make checks payable to: Chief, Photoduplication Service, Library of Congress.

tionship between the second and first variables, on one hand, with the relationship between the second and fourth variables, on the other hand. The simplex pattern requires that the former relationship be higher than the latter one. Now the former two variables differ in the element of Facet B; the latter ones differ in the elements of Facets A and C. Contiguity alone does not tell which relationship will be higher. It may be possible indeed that Facet B will contribute to the relationship more than A and C combined. If, however, one assumes that the various facets have approximately the same influence on the relationship, then counting the common facet elements predicts that the relationship between the former two variables (differing in one facet only) will be higher than the relationship between the latter ones, differing in two facets. This is precisely the prediction required by the simplex pattern.

In a semantic Guttman scale the elements of the facets are related to the order in a monotonic way: As one moves along the order, from the first to the fourth variable, the elements of the facets change progressively, and none of them resumes its former values. First Facet B changes from b_1 to b_2 , then Facet C, and finally Facet A. All the facets behave as the first principal component of the order.

These four variables are a subset of the eight possible variables defined by the three dichotomous facets. The variables not included in the ordered subset are: $a_2b_1c_1$, $a_1b_1c_2$, $a_2b_1c_2$, and $a_2b_2c_1$.

By choosing appropriate variables from the total set of eight variables, it is possible to form other scales, in addition to the one which has been discussed. Thus several subsets are scalable, but the set as a whole cannot

be ordered by the principle of contiguity. For example, Variable $a_1b_1c_1$ has three neighbors, $a_2b_1c_1$, $a_1b_2c_1$, and $a_1b_1c_2$; the principle of contiguity does not say which one of the neighbors is most related to $a_1b_1c_1$.

In a set of eight variables, defined by three dichotomous facets, the principle of contiguity does not suggest an ordering of the whole set; only certain subsets can be ordered. Indeed this principle does not differentiate among the various facets in determining the order. If all the facets play a similar role in the order, the eight variables can be represented as the vertexes of a cube, the dimensions of the cube being the three facets. Three dimensions will be necessary to order the set, as many dimensions as facets. For the same reason, the principle of contiguity is sufficient for proposing a circular order, i.e., a two-dimensional structure, in a set of four variables defined by two dichotomous facets. Let us consider a set of four variables, defined by Facets A and B with Elements a_1 , a_2 , and b_1 , and b_2 , respectively, as follows:

1. a_1b_1
2. a_1b_2
3. a_2b_2
4. a_2b_1

The first variable is contiguous to the second (in Facet A) and to the fourth one (in Facet B), the second variable is contiguous to the first and the third ones, the third variable to the second and the fourth, and the fourth variable to the third and first. Thus, if the principle of contiguity holds, the four variables will go in circle; each variable will correlate higher with its two neighbors and lower with the remaining variable. Examples of such circumplex structures have been reported by Guttman (in

press) in the reanalysis of earlier published correlation matrices of intelligence tests. In these examples Facet A is the type of ability with two elements: a_1 , achievement, and a_2 , analytical ability. Facet B is the language of communication of the test, again with two elements: b_1 , numerical, and b_2 , pictorial. Tests less similar in facet structure correlate less than tests of more similar structure. Thus the circular structure appears.

The contiguity principle proposes the ordering of the variables of the set in as many dimensions as the number of (dichotomous) facets. It does not consider the possibility that a smaller number of dimensions will be sufficient for ordering the variables. Such a possibility implies a differentiation among the facets in determining the order. Some facets will contribute to the determination of the order more than certain other ones. A nonparametric approach to this problem is provided by the new notion of *semantic principal component* (Foa, 1961). The oscillatory functions called principal components are well known to mathematicians and have been used already in the physical and behavioral sciences. Guttman (1954c) identified them empirically in scalable attitudes. Mathematical concepts are formal and, as such, can be employed in a variety of content areas without implying that such different areas have any substantive similarity. The novelty here consists in using the notion of principal component in the analysis of a conceptual structure rather than an empirical one, as done in earlier applications.

SEMANTIC PRINCIPAL COMPONENTS

Let us consider a set of eight variables defined by three dichotomous facets as follows:

1. $a_1b_1c_1$
2. $a_1b_1c_2$
3. $a_1b_2c_2$
4. $a_1b_2c_1$
5. $a_2b_2c_1$
6. $a_2b_2c_2$
7. $a_2b_1c_2$
8. $a_2b_1c_1$

There are several different ways to order these variables in a circle. The problem is how to predict a particular order. Let us suppose, for the sake of argument, that we predict the above order. In this order Facet A behaves as the first semantic principal component; it changes value only once. Facet B behaves as the second component; it changes value twice. Facet C behaves as the fourth component; it changes value four times. The possibility of a third component will be discussed later.

Reversing the argument one can now say that knowledge of the facet components tells a good deal about the order of the variables, but not everything. An alternative order, which preserves the components, may indeed be obtained by interchanging, in the above order, Variable 1 with Variable 2, Variable 3 with Variable 4, and so on, as follows:

2. $a_1b_1c_2$
1. $a_1b_1c_1$
4. $a_1b_2c_1$
3. $a_1b_2c_2$
6. $a_2b_2c_2$
5. $a_2b_2c_1$
8. $a_2b_1c_1$
7. $a_2b_1c_2$

To determine a unique order it is also necessary to identify the first variable of the order. If the first variable is $a_1b_1c_1$, then the alternative order given above is ruled out. Identification of the first variable is also required for differentiating between the first and

the second components. Since the proposed order is circular, one could otherwise start, for example, from the third variable; then Facet A would behave as the second component and Facet B as the first one.

The first variable cannot be identified simply by indicating the subscripts of its facet elements. If these subscripts are assigned arbitrarily, any one of the variables can be indicated by $a_1b_1c_1$ or by any other combination of subscripts we wish. It is therefore necessary to assign subscripts to the elements of the facets in a meaningful manner.

In conclusion: To predict a unique order, for the eight variables defined by three dichotomous facets, it is necessary and sufficient to identify the principal component of each facet and the first variable of the order.

An approach to the solution of these problems has been developed in a study of the facet structure of interpersonal behavior. This study is described in the next section which also suggests a way to determine the principal components of the facets. The problem of identifying the first variable of the order is discussed in the section following the next one.

STRUCTURE OF INTERPERSONAL BEHAVIOR

Interpersonal behavior is defined as the Cartesian product of the *observer* by the *perceptual* and *behavioral* facets.

The perceptual facets are:

Facet A. The person doing the action, or *actor*, with two elements: a_1 , the other (nonobserver), and a_2 , the self (observer).

Facet B. The *level*: b_1 , actual (what is done), and b_2 , ideal (what ought to be done).

Facet C. The person from the point of view of whom the action of a given

actor is perceived, or *alias*: c_1 , the other (nonactor), and c_2 , the self (actor).

The Cartesian product, ABC, of these three perceptual facets defines eight perceptual types, such as actual behavior of Actor₁ from the point of view of Alias₂, and the like.

The three behavioral facets are:

Facet D. *Content* of behavior: d_1 , acceptance or giving, and d_2 , rejection or taking away;

Facet E. *Object* of behavior: e_1 , the other (nonactor), and e_2 , the self (actor);

Facet F. *Mode* of behavior: f_1 , social or status, and f_2 , emotional or love.

The Cartesian product, DEF, of these three behavioral facets defines eight behavioral types, such as emotional acceptance of self, social rejection of other, and the like (Foa, 1961).

The Cartesian space for an observer is defined by the product ABCDEF. Each variable belonging to this space is defined by a profile, over the six facets, of the form $a_1b_1c_1d_1e_1f_1$ ($i,j,k,l,m,n=1,2$). Thus $2^6=64$ variables defined for each observer. The structure of the ABCDEF space as a whole is discussed in the next section. Our problem here is how to predict the order of the eight perceptual types, when the behavioral type is held constant, and the order of the eight behavioral types, when the perceptual type is held constant. That is: the order of the variables ABC ($d_1^6m^6n$) for any given $d_1^6m^6n$ and the order of the variables ($a_1b_1c_1$) DEF for any given $a_1b_1c_1$.

It has been noted already that a necessary condition for predicting the order is to know the component of each facet. It is now proposed that these facets of interpersonal behavior develop at different stages of the process of socialization in the child. Among

perceptual facets it is suggested that differentiation between actors will develop first, followed by differentiation between levels and then between aliases. In the behavioral facets the suggested sequence of development is content, object, mode. Each successive differentiation is obtained by a subdivision of the previous concept according to the new facet (Foa, 1964). In the behavioral facets, for example, the first differentiation is, by content, into acceptance and rejection. At the second stage each one of these concepts splits into two new concepts according to object: acceptance of other and self, rejection of self and other. In the next stage each one of these four concepts is again dichotomized by the mode facet into social and emotional acceptance (or rejection) of self (or other). This three-stage pattern of dichotomization suggests a circular order of the variables in which the first facet of the sequence behaves as the first principal component, the second facet as the second component, and the remaining facet as the fourth component (see Footnote 2). The sequence of concept differentiation by facets corresponds to the order of the components. Since the process of concept differentiation is symmetric, the third component does not appear. Odd components are antisymmetric; they have a different facet element in the first and last variable. Even components, on the other hand, are symmetric. More generally, it may be suggested tentatively that, when the predicted order is a *circumplex* (a circular order of correlation coefficients), even components are more likely to appear since the first and last variables are also neighbors. When the order is a simplex, odd components may be more likely. Two examples of simplexes, in sets of eight variables, in which one of the three facets behaves as third component of the order, have

been found recently by H. Triandis of the University of Illinois.³

Let us now go back to our study of interpersonal behavior. Using the components of the facets, as determined by the sequence of development, we can predict that the circular order of the perceptual types may be $a_1b_1c_1$, $a_1b_1c_2$, $a_1b_2c_2$, $a_1b_2c_1$, $a_2b_2c_1$, $a_2b_2c_2$, $a_2b_1c_2$, $a_2b_1c_1$, and the circular order of the behavioral types, $d_1e_1f_1$, $d_1e_1f_2$, $d_1e_2f_2$, $d_1e_2f_1$, $d_2e_2f_1$, $d_2e_2f_2$, $d_2e_1f_2$, $d_2e_1f_1$.

The Cartesian product of the 8 perceptual types \times the 8 behavioral types defines 64 variables of interpersonal behavior as perceived by a given observer. To study the intercorrelation among perceptual types, we shall choose sets of 8 variables having the same behavioral type and observer and differing in perceptual type. Likewise, to study the intercorrelation among behavioral types, we shall keep the perceptual type and observer constant and let the behavioral type vary. These variables were observed in a sample of 633 married couples in Jerusalem, Israel. The procedure for gathering the data has been described elsewhere (Foa, 1962). Here it may suffice to say that the frequency of perceived occurrence of each one of the variables was recorded as reported by the observer (husband or wife) by means of a three-question scale. The scale score on each variable was then intercorrelated with the score of other variables. Sixteen intercorrelation matrices were computed for the perceptual types, holding the behavioral type and the observer constant. Likewise 16 intercorrelation matrices were computed for the behavioral type, holding the perceptual type and the observer constant (Foa, 1962).

³ H. Triandis, personal communication, 1964.

In all these matrices the predicted circumplex order is confirmed. The coefficients are high near the main diagonal; as one moves away from it the coefficients tend to decrease and then to increase again as the main diagonal is approached from the other side (see Footnote 2).

These examples suggest that if we are able to make assumptions about the sequence of development of facets we may also be able to propose hypotheses about the order of variables defined by these facets: earlier developing facets will behave as lower components.

It has been noted already that knowledge of the facet components is not sufficient for determining a unique order. To this end it is also necessary to identify the first variable of the order. Some considerations leading to a possible solution of this problem will be introduced in the next section.

SPECIFICITY OF THE FACET ELEMENTS

So far we have been concerned with the relationship between a variable and every other variable belonging to the set. Let us now consider the relationship between a variable and the set as a whole. The set may be designated by a facet which is constant and common to all the variables belonging to it; in our example this facet is interpersonal behavior.

The set "interpersonal behavior" can be seen as a subset of a larger set which may be called "behavior." Another subset of the behavior set is "*personal behavior*," which may be defined as the behavior which does not require the actual or potential participation of more than one person in order to occur.

In some of the facets of interpersonal behavior one element characterizes variables which belong to the interpersonal set only, while the other element indicates variables which belong also

to the personal set. Thus in the actor, alias, and object facets the element *other* defines variables belonging to the interpersonal set only, while the element *self* defines variables which also belong to the personal set.

Likewise, the element *social* of the facet mode defines variables belonging to the interpersonal set only, while the element *emotional* indicates variables belonging to both sets.

A partition of the interpersonal set is suggested by the elements of the remaining two facets, content and level. Variables with elements *acceptance* and *actual* must occur in interpersonal behavior, while variables with the elements *rejection* and *ideal* may or may not occur. Rejection and ideal behavior are not necessary for the occurrence of interpersonal behavior.

Let us call a facet element which belongs to the interpersonal set only, or which must occur in interpersonal behavior, a *specific* element.

Thus the elements *other*, *social*, *acceptance*, and *actual* are specific to interpersonal behavior; while the elements *self*, *emotional*, *rejection*, and *ideal* are not specific.

Less technically, the notion of specificity can be described in the following manner: Relations with *other* must occur in interpersonal behavior, obviously, while relations within *self* need not. Social behavior is obviously interpersonal while emotional behavior can occur without reference to *other*. Acceptance must occur in interpersonal behavior if the behavior is to continue to be interpersonal, while rejection leads to cessation of interpersonal behavior. Actual behavior is required in an interpersonal situation while the concept of the ideal is not and can be maintained in a covert cognitive field.

In each of these comparisons the element associated with interpersonal behavior is labeled "specific." It is

now proposed that this classification of the elements of the facets is helpful in predicting the multiple correlation between a variable and the other variables of the set. Given two variables, one including the specific element and the other one the nonspecific element of a given facet, and with the same element composition in the other facets, the multiple correlation of the variable with the specific element will be higher than the multiple correlation of the variable with the nonspecific element. The findings support this hypothesis (Foa, 1963). For example, when the element acceptance appears in the variable its multiple correlation is always higher than the corresponding multiple correlation of the variable in which the element rejection appears. In the other facets there are some deviations from the prediction, but the hypothesis is significantly sustained by the results of the sign test (see Footnote 2).

This approach to the analysis of the relationship between a variable belonging to a set and the set as a whole suggests that some facet elements are specific to this set while some others are not. Variables containing the specific element are more strongly related to the set than those which do not contain it.

The nonspecific elements may be seen as a link between a particular area and other neighboring areas. As such they may prove helpful in clarifying the relationship between different areas, leading, ultimately, to more general theoretical constructs.

In each facet one element is specific, while the other is nonspecific. Thus the two elements can be ordered according to their specificity. Let us assign the subscript 1 to the specific element and the subscript 2 to the nonspecific one. This convention has been followed in defining the facet structure of interpersonal behavior. In

this manner a meaning has been given to the subscript; it indicates the specificity of the element. It is now possible to obtain a unique order of the variables determined by the principal component of the facets. To this end it is sufficient to start the order from the variable containing only specific elements, i.e., from Variable $a_1b_1c_1$ in our notation, and to continue with the other variables according to the principal components of the facets. This is precisely what has been done in proposing an order for the perceptual and behavioral types of interpersonal behavior. The alternative order would be obtained by starting with Variable $a_1b_1c_2$. Further study is required to find out when to predict one order or the other one. For this particular case of interpersonal behavior the meaning of the order which has been chosen is quite clear. It just proposes that, in the perceptual types, the ideal and actual perceptions will be more related for the alias of the actor than for the alias of the nonactor. In the behavioral types the behavior toward the self and toward the other will be more related in the emotional than in the social mode. More generally, given the four variables $a_1b_1c_1$, $a_1b_1c_2$, $a_1b_2c_2$, and $a_1b_2c_1$ two variables differing in the element of B and alike in C will be more related when the element of C is nonspecific, c_2 , than when it is specific, c_1 . That is, $a_1b_1c_2$ and $a_1b_2c_2$ are closer than $a_1b_1c_1$ and $a_1b_2c_1$.

Specificity provides a common ground for relating the components of the facets to the order of the variables. Moving along, the order the specificity of the first facet changes only once. In the second facet it changes twice and in the last facet four times (see Footnote 2).

Another application of the notion of specificity is in predicting the empirical structure of the ABCDEF set as a

whole. This notion, indeed, lends itself easily to a structural interpretation. Specific elements belong to the ABCDEF space only; it seems, therefore, that they will be found in the variables situated in the central portion of this space. Nonspecific elements, on the other hand, belong to other spaces as well and are therefore likely to appear in the variables situated on the periphery of the ABCDEF space, nearer to other spaces. This structural interpretation of specificity is supported by the finding that specific variables are related to the set as a whole, more than nonspecific ones. It seems therefore that the ABCDEF structure will be such that specific variables will occupy a central position with respect to nonspecific ones.

It has been seen earlier that every perceptual subset of eight variables of the form $ABC(d_1e_mf_n)$, and every behavioral subset, also of eight variables of the form $(a_1b_jc_k)DEF$, constitute an empirical circumplex. It follows that the whole ABCDEF set is a net of intersecting circumplexes forming either a hyperspheroidal surface or an anchor-ring surface. The anchor ring or torus is a figure shaped like the inner tube of a car.

In a spheroidal surface the distinction between center and periphery does not apply: None of the points on such a surface seems to be more central than any other. This configuration might be appropriate when facet elements cannot be separated into specific and nonspecific ones. When this distinction is possible, as in our case, the predicted structure is the anchor ring. In this figure variables on the inner part of the surface (where the air valve is found in a car tube) are more central than those situated on the outer part of the surface. We propose to call a statistical structure of variables arranged on the surface of an anchor

ring a *ringex*. Thus the notion of specificity, and its structural interpretation as centrality, leads to the prediction of a ringex structure for the ABCDEF space.

In this ringex the perceptual subsets of the form $ABC(d_1e_mf_n)$ constitute small circles on the section of the torus, while the behavioral subsets $(a_1b_jc_k)DEF$ constitute large circles along the other direction of surface, across the perceptual circles.

Each perceptual circle crosses each perceptual circle once, and this point of intersection locates a variable of the form $a_1b_jc_kd_1e_mf_n$.

The ringex hypothesis suggests that the perceptual circles will be smaller than the behavioral circles. Among behavioral circles those situated on the inner portion of the surface will be the smallest, and the size of these circles will increase as one moves from the inner to the outer portion. To test these properties of the ringex the average correlation among variables belonging to each given circle was computed. This correlation may be interpreted as inversely related to the size of the circle. The results fully support the hypothesis. In particular, the average correlation of the behavioral circumplexes $(a_1b_jc_k)DEF$ is highest when the elements $a_1b_jc_k$ are specific and becomes progressively lower as the elements change to nonspecific.

Specific variables are more central than nonspecific ones, and the distance among variables increases as one moves from the inner to the outer portion of the ringex.

The notions of sequence of development of facets and of specificity of facet elements have led to the prediction of certain empirical structures, the circumplex and the ringex, respectively. When the facet design is such that these notions do not apply, the empirical prediction may be different.

If the facets do not develop in a sequential manner, an ABC design may produce a cubex structure rather than a circumplex. If specificity does not apply, we may have a spherex instead of a ringex. Thus these notions may provide criteria for differentiating among different empirical structures resulting from formally identical facet designs.

FACETS AND CATEGORY FORMATION

In the examples of studies described here facet design has led to the correct prediction of empirical result. It would be possible to quote other examples in which this did not happen. Systematic design alone does not guarantee that the data will sustain the hypothesis. This is hardly surprising in view of the fact that alternative designs are usually possible, leading to different hypotheses.

It may be suggested that correct prediction is more likely to occur when there is a certain correspondence between the design devised by the investigator and the manner in which the universe of experience being investigated is categorized by the subject. If psychological equivalence is a result of the coding operations of the subject rather than a property of the stimulus, then the problem of inventing a satisfactory facet design appears to be closely related to the psychology of concept formation. It becomes relevant to ask with Bruner (1956, p. 8) how different conditions, situational, cultural, in the personal history of the subject may lead to a given categorization rather than to another one. Some of the experiments described by Bruner for investigating problems of concept formation have a straight facet design. Several experiments, for example, are based on a Cartesian product of four facets, each having three elements (p. 42).

Bruner recognizes the influence of language in category formation and suggests that individual differences in categorizing may be found in affective responses developed in early childhood before the acquisition of language.

On the other hand, when the language provides a name for labeling a facet, this fact in itself may indicate that such a particular manner of classifying experience has gained acceptance beyond the experienced world of a given individual. Nevertheless, the development of language may blur the original pattern of concept formation. Zipf (1935) has shown that in several languages the length of a word tends to be inversely related to the frequency of its usage. When the "automobile" becomes the "car," the identity of the component elements is lost. This may be one of the reasons why, in Indo-European languages, the sound used to designate a concept rarely suggests the component elements of the concept. Exceptions are such multisyllabic words as *chairman*, *newspaper*, and *underwear*. In other languages, such as Chinese, new concepts may be indicated by combining the sounds of the component elements, which is equivalent to using multisyllabic words whose total meaning is a function of the meaning of their syllables. This latter type of language may present fewer obstacles in laying bare the mechanism of concept formation than do Indo-European languages.

One may suggest that the unificatory role of language may be lost in cross-cultural research. Yet it appears that a concept found in one language may usually be rendered in another one by a phrase, if not by a word (Brown, 1956). It is therefore possible that certain basic facets may be common to different cultures. Some evidence is available on this point.

The study of intergroup beliefs and action which has been summarized in this paper was originally done in Brazil. The same design was later used in France with similar results (Foa & Guttman, 1962).

In studying the structure of interpersonal behavior it has been shown that the same order of intercorrelation appears in respondents with a European and a Middle Eastern background, although the size of the correlation may change (Foa, 1964).

As we gain more knowledge in the process of concept attainment and in the role played by language in this process, we may be better able to produce satisfactory facet designs. On the other hand, facet design and analysis may prove useful in advancing knowledge in this area, as well as in other behavioral investigations. An assessment of the value of facet design and analysis, as an approach to multivariate research, will require more applications. The purpose of this paper has been to give a view of the present situation and to suggest that the results obtained so far may warrant its further use.

REFERENCES

- BROWN, R. W. Language and categories. In J. S. Bruner, Jacqueline J. Goodnow, & G. A. Austin (Eds.), *A study of thinking*. New York: Wiley, 1956. Pp. 247-312.
- BRUNER, J. S., GOODNOW, JACQUELINE J., & AUSTIN, G. A. (Eds.), *A study of thinking*. New York: Wiley, 1956.
- BRUNSWIK, E. Note on Hammon's analogy between "Relativity and representativeness." *Philosophy of Science*, 1951, 18, 212-217.
- BRUNSWIK, E. *Perception and the representative design of psychological experiments*. Berkeley: Univer. California Press, 1956.
- FISHER, R. A. *The design of experiments*. (5th ed.) London: Oliver & Boyd, 1949.
- FOA, U. G. The contiguity principle in the structure of interpersonal relations. *Human Relations*, 1958, 11, 229-238.
- FOA, U. G. Convergences in the analysis of the structure in interpersonal behavior. *Psychological Review*, 1961, 68, 341-353.
- FOA, U. G. The structure of interpersonal behavior in the dyad. In Joan Criswell, H. Solomon, & P. Suppes (Eds.), *Mathematical methods in small group processes*. Stanford: Stanford Univer. Press, 1962. Pp. 166-179.
- FOA, U. G. A facet approach to the prediction of communalities. *Behavioral Sciences*, 1963, 8, 220-226.
- FOA, U. G. Cross-cultural similarity and difference in interpersonal behavior. *Journal of Abnormal and Social Psychology*, 1964, 68, 517-522.
- FOA, U. G., & GUTTMAN, L. Facet design and analysis of data on personality and attitudes related to human organization. Technical Final Report, March 1962, Contract No. A. F. 61(052)-121, Israel Institute of Applied Social Research. (Mimeo)
- GUTTMAN, L. A new approach to factor analysis: The radex. In P. R. Lazarsfeld (Ed.), *Mathematical thinking in the social sciences*. Glencoe, Ill.: Free Press, 1954. (a)
- GUTTMAN, L. An outline of some new methodology for social research. *Public Opinion Quarterly*, 1954, 18, 395-404. (b)
- GUTTMAN, L. The principal components of scalable attitudes. In P. F. Lazarsfeld (Ed.), *Mathematical thinking in the social sciences*. Glencoe, Ill.: Free Press, 1954. (c)
- GUTTMAN, L. A structural theory for intergroup beliefs and action. *American Sociological Review*, 1959, 24, 318-328.
- GUTTMAN, L. A faceted definition of intelligence. In *Scripta Hierosolymitana*. Jerusalem: Hebrew Univer., in press.
- HAMMOND, K. R. Relativity and representativeness. *Philosophy of Science*, 1951, 18, 208-211.
- STEPHENSON, W. *The study of behavior*. Chicago: Univer. Chicago Press, 1953.
- ZIPP, G. K. *The psycho-biology of language*. Boston: Houghton Mifflin, 1935.

(Received February 18, 1964)

THE MOTOR THEORY OF SPEECH PERCEPTION: A CRITICAL REVIEW¹

HARLAN LANE

Center for Research on Language and Language Behavior, University of Michigan

The motor theory of speech perception maintains that articulatory movements and their sensory feedback mediate between the acoustic stimulus and the perception of speech. The theory is based on examination of changes in identification probability, identification latency, and discrimination accuracy effected by changes in synthetic speech stimuli. This paper reviews first those experiments cited in support of the theory, then opposing evidence is presented: It is shown that identification and discrimination functions for nonspeech stimuli do not differ from those for speech stimuli, when obtained under comparable conditions.

The motor theory of speech perception maintains that "articulatory movements and their sensory effects mediate between the acoustic stimulus and the event we call perception [Liberman, 1957, p. 122]." The observed sequence of events in a speech-perception episode may be represented as:

$$R_V \rightarrow S_A \rightarrow R_D;$$

a vocal response, R_V , generates an acoustic stimulus, S_A , which leads to a discriminative response, R_D . According to the motor theory of speech perception, the expanded sequence of

events is:

$$\begin{array}{ccc} R_V & \rightarrow & S_A \quad R_D \\ & \downarrow & \uparrow \\ & R_V' & \rightarrow S_P \end{array};$$

the acoustic stimulus leads to a covert, articulatory response, R_V' , whose proprioceptive feedback, S_P , leads to the discriminative response.

The inference of mediating articulation in speech perception has appeared in several experimental contexts, including discussions of phonetics and phonemics, the discrimination of synthetic speech stimuli, the electromyography and acoustics of speech production, psychophysical functions for speech parameters, automatic speech recognition, disorders of speech and hearing, first-language acquisition, and second-language pedagogy.² Although

¹ This investigation was supported in part by the Language Development Section, United States Office of Education (Contract Number SAE-9265), and by the National Science Foundation (Grant Number 05586). D. V. Cross and B. A. Schneider have collaborated in the conduct of the research and its description in this article. I have also profited by the comments of A. M. Liberman, Haskins Laboratories, who read an earlier draft and first introduced me to the analysis of speech discrimination. But I am chiefly indebted to N. J. Peterson, Yale University, for many suggestions, criticisms, and corrections.

² In addition to those articles whose findings directly concern the validity of the mediation hypothesis and are discussed in this review, the following articles have cited the hypothesis in discussions of theory or experimental results: Stevens (1960): "Thus, in the synthesis process . . . a representation of the signal at the articulatory level will certainly occur . . . a similar representation may likewise exist at some

the hypothesis that speech is perceived by reference to articulation has not been elaborated into a formal theory, its recurrence and central role in these discussions suggest that a critical re-

stage during the reverse process of speech recognition [p. 53]"; Fant (1963): "To me the reference to articulation primarily serves a function within the metalanguage whereby we as outside observers may conveniently describe speech. But is it actually a part of speech perception? . . . The alternative view I would like to propose here is that if the auditory analysis in the hearing process has proceeded so far as to allow the proposed articulatory matching, the decoding could proceed without an articulatory reference [p. 1]"; Jakobson and Halle (1956): "The theoretically unlikely surmise of a closer relationship between perception and articulation than between perception and its immediate stimulus finds no corroboration in experience: the kinaesthetic feedback of the listener plays a very subordinate and incidental role. Not seldom do we acquire the ability to discern foreign phonemes by ear without having mastered their production . . . [p. 34]"; Hockett (1955): "Similarly, we may suspect that [as Jack] listens to Jill, his Speech Receiver is able to decode the signal partly because the incoming signal is constantly compared with the articulatory motions which Jack himself would have to make in order to produce an acoustically comparable signal . . . in learning a foreign language, one has considerable difficulty hearing correctly until one can also pronounce correctly [p. 7]"; Prins (1963): "Children who tended to confuse place of articulation during speech sound production also had difficulty discriminating minimal word pairs in which a single phoneme was altered in place of articulation. This finding when related to previous psycholinguistic research suggests that sound discrimination ability is a function of articulation . . . [p. 387]"; Lenneberg (1962): "A case [has been] presented, typical of a larger category of patients, where an organic defect prevented the acquisition of the motor skill necessary for *speaking a language* but evidence was presented for the acquisition of grammatical skills as required for a complete understanding of language [p. 424]"; also see Chistovich, Klaas, and Alekin (1961, p. 173); Delattre (1958, p. 228); Fonagy (1958, p. 58); Lane (1965, p. 32); and Thurlow (1963, p. 469).

view is warranted. The purpose of this article is to bring together and examine the various experimental findings that have been interpreted as favoring a motor theory of speech perception, to present opposing findings and interpretations, and to show that, in the light of the available evidence, the postulation of a special perceptual mechanism for the discrimination of speech is unjustified.

The research findings that have been interpreted as favoring this theory have in common that the overt verbal behavior of the listener (R_D) in response to speech is more closely correlated with the patterns of articulation (R_V) that typically produce those speech signals than with the acoustic signals (S_A) themselves. From this observation it is inferred that corresponding patterns of articulation (R_V') actually mediate between the acoustic signals and the overt reports of the listener.³ Since the speech signals generated by articulation are predictable from anatomical data (Fant, 1960), a correlation between the perceptual response and the acoustic stimulus necessarily implies a correlation between the perceptual response and the pattern of articulation that generated the stimulus. Therefore, the mere correlation of perception with articulation does not warrant the introduction of a mediating process in analyzing the listener's behavior in response to speech signals. In order to establish the usefulness of this inference, the experimental evidence must show that

³ In the most recent explication of the theory, Liberman, Cooper, Harris, and MacNeilage (1963) suggest that the mediating events may be one or two steps removed from articulation: "The neural commands themselves, or rather their equivalents in the central nervous system, might be used to provide the reference system in terms of which the decoding [of speech] is carried out [p. 8]."

the perceived similarities and differences of a set of speech sounds correspond *more closely* to the articulatory than to the acoustic similarities and differences among the sounds. In other words, R_V' is inferred when there is a greater isomorphism between changes in R_V and R_D than between S_A and R_D .⁴

There are fundamentally three kinds of evidence, therefore, that have been viewed as favoring a motor theory of speech perception; these serve as the framework for the present discussion:

⁴In linguistic terms, Liberman, Cooper, Harris, and MacNeilage (1963) have put it this way: "The articulatory reference theory [rests] empirically on the evidence which indicates that the relation between phoneme and articulation is more nearly one-to-one than is the relation between phoneme and acoustic signal [p. 7]."

(a) when diverse perceptual responses are evoked by a set of acoustically similar speech signals produced by diverse patterns of articulation, (b) when similar perceptual responses are evoked by a set of acoustically diverse speech signals produced by similar patterns of articulation, and (c) when both articulation and the acoustic signal change continuously over some range, and changes in the perceptual response correspond more closely to changes in articulation.

EVIDENCE THAT FAVORS A MOTOR THEORY OF SPEECH PERCEPTION

Similar Acoustic Stimuli, Diverse Perceptual and Articulatory Responses

Over the past few decades, scientists at the Haskins Laboratories have greatly advanced our understanding

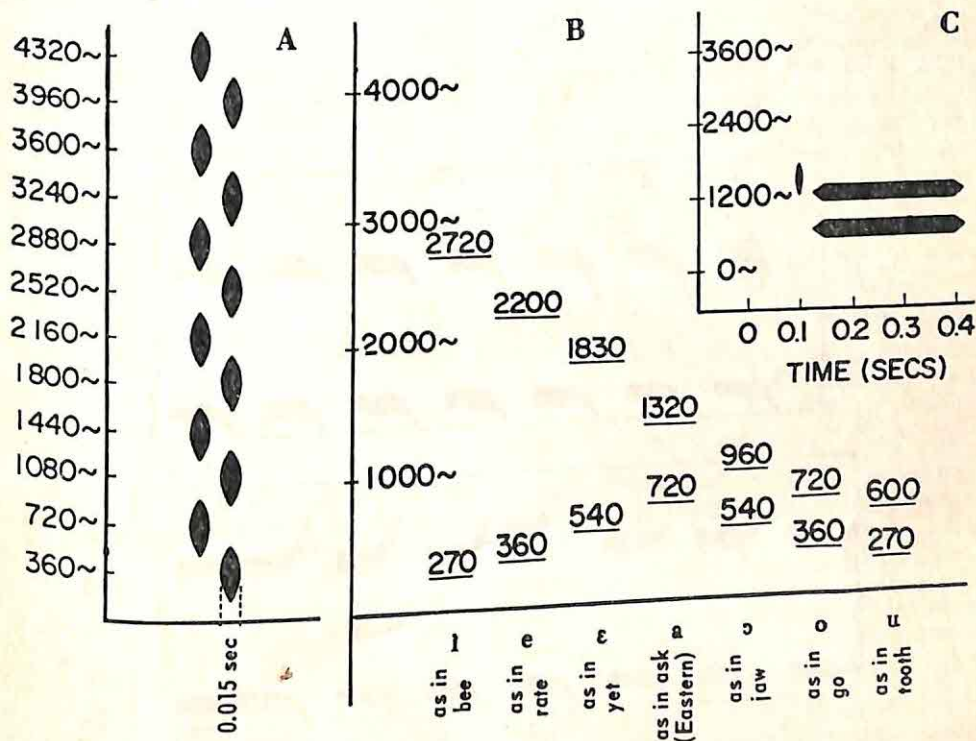


FIG. 1. The frequency positions of the schematic stops (A) and vowel formants (B) that were combined (as in C) and converted into sound for identification by listeners. (After Liberman et al., 1952.)

of the perception of speech by experiments involving synthetic speech stimuli, generated by a device called Pattern Playback, which permits careful control of stimulus parameters.⁵ Among the earliest studies whose findings are seen as favoring a motor theory of speech perception is an investigation of the perception of unvoiced stop consonants by Liberman, Delattre, and Cooper (1952), employing the Pattern Playback to convert hand-painted spectrograms into sound. Figure 1, Part C, shows one of the schematic syllables that was converted into sound, and Figure 1, Parts A and B, shows the two acoustic variables of the experiment: the frequency locus of a schematic stop and the formant frequencies of the following vowel. The 12 stops before each of the 7 vowels gave a set of 84 consonant-vowel (CV)

⁵ For a description of Pattern Playback and its principles of operation, see Cooper (1950, 1953) and Cooper, Liberman, and Borst (1951).

syllables that were presented to subjects for identification.

When the frequency locus of the schematic stop was held constant and the following vowel was varied over the seven pairs of formant frequencies, shifts in the modal consonant identification were obtained. For example, the schematic stop at 1,440 cycles per second was called /p/ (59 times out of 60) before /i/, but it was increasingly often called /k/ as the following vowel ranged from /i/ to /a/; it was called /k/ (57/60) before /a/ but increasingly often called /p/ as the following vowel ranged from /a/ to /u/; the same stop was called /p/ (42/50) before /u/.

In order to account for the diverse consonant identifications evoked by a set of invariant schematic stops, the investigators pointed to the diverse patterns of articulation necessary to generate the natural-language correlates of their acoustic stimuli:

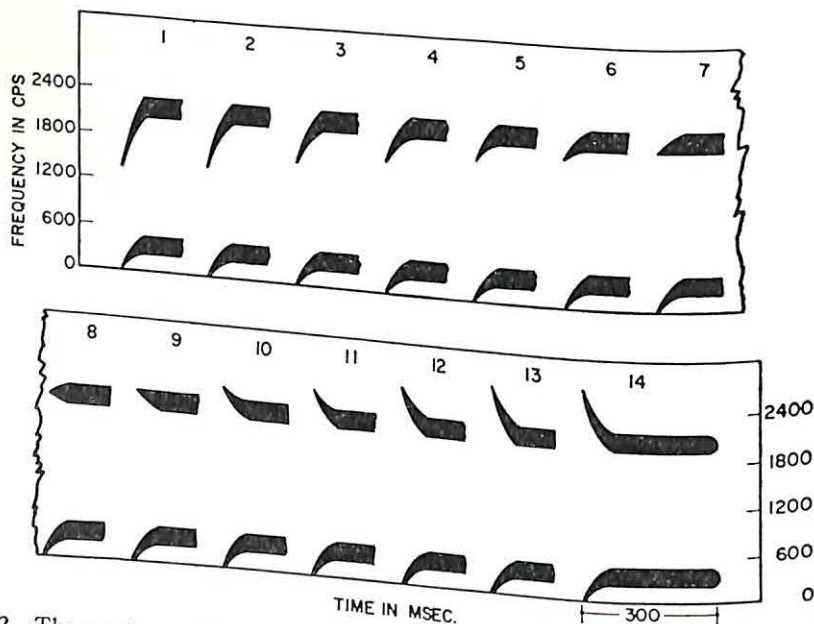


FIG. 2. The spectrographic patterns, varying in the direction and extent of second-formant transition, that were converted into sound for identification as /b/, /d/, or /g/. (From Liberman et al., 1957.)

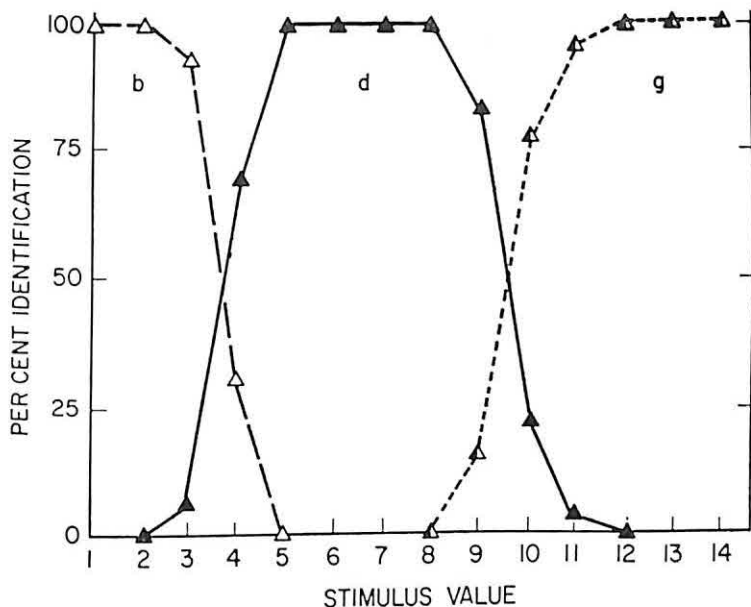


FIG. 3. Identification of the consonants synthesized with the spectrographic patterns shown in Figure 2. (Each point shows the relative frequency of the responses /b/, /d/, and /g/ in 32 presentations of each stimulus to one listener. After Liberman et al., 1957.)

With the mouth set to articulate /i/ or /u/, a speaker would presumably use a movement of the lips to produce a sound having roughly the same acoustic characteristics as our schematic stop at 1440 cps; for the vowel /a/, the same acoustic result would be most closely approximated by lowering the back of the tongue from a point of contact with the roof of the mouth [Liberman et al., 1952, p. 513].

The authors then suggest that these diverse patterns of articulation are operative during acoustic reception to translate acoustic similarity into perceptual diversity:

We assume, then, that the sound patterns corresponding to the syllables *pi* (or *pu*) and *ka* set off the appropriate movements in the listener (that is, the articulatory movements which the listener would make in attempting to reproduce these acoustic patterns) or perhaps only the short-circuited neural equivalents of these movements, with the result that the initial schematic phonemes *p* and *k*, which can be entirely identical as acoustic stimuli, become clearly differentiated in proprioceptive terms and are therefore finally perceived as distinctly different sounds [p. 514].

Several more recent experiments by Liberman and co-workers have also examined labeling responses to synthetic consonant-vowel pairs, and have provided support for the motor theory of speech perception. The five studies summarized below have in common that the acoustic parameters of the stimuli that controlled consonant identification all varied regularly over some range, while the parameters of the following vowel were held constant.

Liberman, Harris, Hoffman, and Griffith (1957) synthesized a series of 14 speech stimuli, using the patterns shown in Figure 2, which varied only in the direction and extent of second-formant transition. When the auditory stimuli were presented in irregular order to a group of listeners for identification, the labeling functions that were obtained resembled the one shown in Figure 3 for a typical subject. The relatively steep gradients that define the phoneme boundaries /b-/d/ and

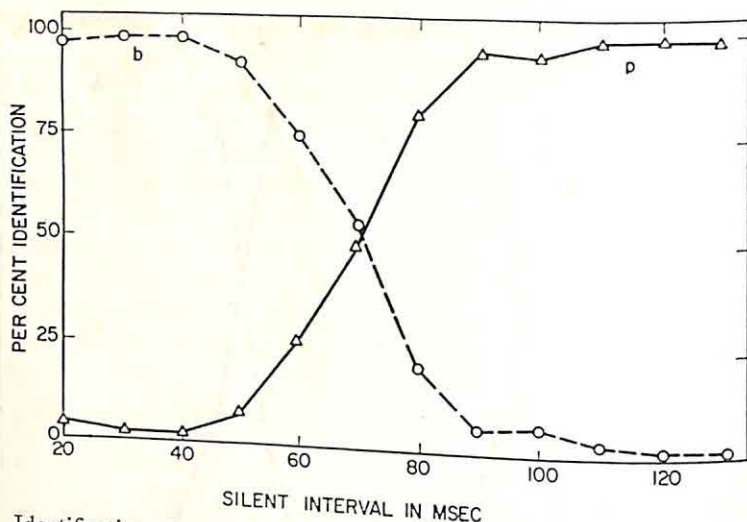


FIG. 4. Identification of synthetic speech stimuli as /ræbɪd/ or /ræpɪd/ as a function of the duration of the silent interval between syllables. (Each point shows the mean relative frequency of the responses /b/ and /p/ in 32 presentations of each stimulus to 12 listeners. From Liberman, Harris, Eimas, Lisker, & Bastian, 1961.)

/d/-/g/ are seen by the authors as favoring the inference of mediating articulation. The articulatory patterns that, in natural language, generate the three phonemes may be viewed as diverse (i.e., discretely different): /b/ is a bilabial plosive, /d/ is alveolar, and /g/, velar. The gradients, which "partition" the stimulus continuum, reflect diverse labeling responses. However, the speech stimuli, drawn from an acoustic continuum, are considered similar. It is inferred, therefore, that diverse mediating responses "map" the stimulus continuum onto diverse perceptual responses.

A later study by Liberman, Harris, Kinney, and Lane (1961) provides additional evidence of this kind for the motor theory of speech perception. Seven synthetic speech stimuli were employed that varied only in the delay of the onset of the first formant relative to that of the second and third formants and that ranged perceptually from /do/ to /to/. The "cutback" in the first formant ranged from 0 to 60 milliseconds in 10-millisecond steps.

The labeling function for a typical subject shows greater than 95% /do/ identification for Stimuli 1 and 2 (= 5% /to/), 80% at Stimulus 3, 14% at Stimulus 4, 5% at Stimulus 5, and no /do/ identifications at Stimuli 6 and 7. "For all Ss, the phoneme boundary between /d/ and /t/ is quite sharp; in every case a change of 10 msec. in the first-formant cutback is sufficient to shift the responses from 75% /d/ to 75% /t/ [p. 383]." The near dichotomy in labeling responses corresponds to the voiced-voiceless dichotomy in classifying the articulations that produce /d/-/t/ in natural language.

The next study in this series, conducted by Liberman, Harris, Eimas, Lisker, and Bastian (1961), employed the phonemic contrast /ræbɪd/-/ræpɪd/. Twelve speech stimuli were synthesized with intervocalic silent intervals that varied from 20 to 140 milliseconds in steps of 10 milliseconds. Figure 4 shows the relative frequency of /b/ and /p/ responses to these stimuli by 12 listeners. The investigators

conclude from these data that, once again, "the phoneme boundary is reasonably sharp [p. 183]." As in the prior study, the corresponding dichotomy in classifying articulation of the speech stimuli in natural language is voiced-voiceless.

The next pair of studies provides additional evidence of "categorical perception" of speech stimuli and also electromyographic data to confirm that the patterns of articulation that generate these stimuli are similarly categorical. Bastian, Eimas, and Liberman (1961) inserted silent intervals of varying lengths into natural recordings of the word /slit/ and presented them to listeners for identification as either /slit/ or /split/. "The perception of this continuum can be regarded as essentially categorical. . . . The categorical perception of the consonants may be explicable in terms of the categorical nature of their articulatory gestures [i.e., the absence or presence of the bilabial plosive; p. 842]."

Harris, Bastian, and Liberman (1961) have shown that, when subjects were instructed to mimic stimuli sampled from the /slit/-/split/ continuum, the silent intervals in their responses fell into a bimodal frequency distribution; they were, in other words, similarly categorical. To determine whether there were incipient /p/ movements in the region of the phoneme boundary, electromyographic recordings were made during mimicry; these were also entirely categorical: "Either there was a normal burst of muscle potential at the lip (indicating a /p/ gesture) or there was not [p. 842]."

In each of the phoneme-labeling experiments described above, the investigators also presented their listeners with the synthetic speech stimuli arranged in ABX triads, in order to measure percent-correct discrimination within pairs of stimuli drawn from the

speech continuum. In accordance with the motor theory of speech perception, the experimenters expected to find that two stimuli drawn from the same phoneme category (defined by the labeling gradients) are less often correctly discriminated than two stimuli drawn from opposite sides of the phoneme boundary. The rationale is spelled out by Liberman, Harris, Eimas, Lisker, and Bastian (1961):

We believe that in the course of his long experience with language a speaker (and listener) learns to connect speech sounds with their appropriate articulations. In time, these articulatory movements and their sensory feedback (or, more likely, the corresponding neurological processes) become part of the perceiving process, mediating between the acoustic stimulus and its ultimate perception. When significant acoustic cues that occupy different positions along a single continuum are produced by essentially discontinuous articulations . . . the perception becomes discontinuous (i.e., categorical), and discrimination peaks develop at the phoneme boundary [p. 177].

The experiments to be described, which examined ABX discrimination of synthetic speech stimuli, had two further objectives. Their second objective was to see to what degree the data warrant the extreme assumption that "the listener cannot hear any differences among these sounds beyond those that are revealed by his use of the phonemic labels [Liberman et al., 1952, p. 359]." This outcome may be expected if both the labeling responses and ABX discriminations are entirely dependent on the proprioceptive feedback from the mediating articulation hypothesized by the motor theory of speech perception. The contrary assumption, namely that the listener can discriminate among many more stimuli than he can identify absolutely, is in accord with psychophysical data on the perception of nonspeech stimuli (e.g., Miller, 1956).

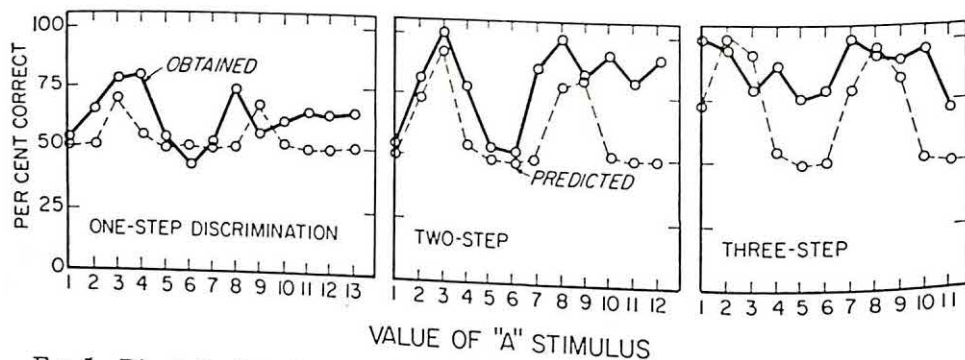


Fig. 5. Discrimination of the synthetic speech stimuli shown in Figure 2, when presented in ABX triads to one listener. (The predicted functions are based on the labeling gradients for this subject, shown in Figure 3. Each point on the obtained functions shows the relative frequency of correct discriminations in 42 presentations of each triad. After Liberman et al., 1957.)

The third and last objective of the discrimination studies that bears on the motor theory of speech perception was to assess the degree to which the discrimination functions reflected "acquired similarity," "acquired distinctiveness," or both:

One possibility [acquired similarity] is that in the raw these speech stimuli were all as highly discriminable as the most discriminable pair, and that the effect of our linguistic experience has been to dull our sensitivity to the differences within the phoneme category . . . the contrary possibility [acquired distinctiveness] is that our stimuli were originally as little discriminable as the least discriminable pair, and that our discriminations have been selectively sharpened as a result of our long experience with the language [Liberman, 1957, p. 122].

The motor theory of speech perception is seen by Liberman as favoring the latter outcome:

It is tempting to speculate that what is learned is simply a connection between various acoustic stimuli and certain articulatory responses. . . . We should suppose that the perceived difference between two relatively similar external stimuli could be increased if we could attach to those stimuli two very different mediating responses and hence gain the added distinctiveness of their very different proprioceptive returns [p. 123].

Discriminations among synthetic speech stimuli were measured, in the

first of the phoneme-labeling experiments described above (Liberman et al., 1957), by arranging pairs of the 14 stimuli in triads of the form ABA and ABB. In this "ABX" procedure, the first stimulus in a triad, A, is the i^{th} stimulus in the series of 14 speech sounds, the second stimulus, B, is the $i + 1^{\text{th}}$ or $i + 2^{\text{th}}$ or $i + 3^{\text{th}}$ (one-, two- and three-step comparisons, respectively), and the third stimulus, X, is either the A stimulus or the B stimulus—the subject is required to say which. The discrimination functions obtained for a typical listener are shown in Figure 5. Comparison of these data with the corresponding labeling gradients in Figure 3 shows that the listener discriminates more accurately between sounds that lie on opposite sides of the phoneme boundary than he does between sounds (separated by the same number of physical units) drawn from within one phoneme category.

In order to assess the degree to which the listener "cannot hear differences . . . beyond those revealed by the phoneme labels," a set of predicted discrimination functions was derived from the labeling gradients, on the basis of three assumptions:

1. Two speech stimuli are discriminated only to the extent that they are labeled differently (assigned to different phoneme classes).

2. The phoneme classes to which the stimuli are assigned during the labeling task are the same as those to which they are assigned (by mediating processes) during the discrimination task, and the relative frequencies of assignment are the same in the two cases.

3. The probability of assigning the i^{th} stimulus in a discrimination triad to the j^{th} phoneme class is independent of the assignments for the other stimuli in the triad.

Figure 5 shows that, in the case of the /b/-/d/-/g/ continuum, the discrimination functions predicted from the labeling gradients correspond in form to the obtained functions but lie somewhat below them. The authors conclude that listeners discriminate the speech stimuli largely to the extent that they label them differently, but they are able to extract some information in addition to that reflected in the labeling gradients.

In order to assess the degree to which the speech-discrimination functions reflect acquired similarity, acquired distinctiveness, or both, it is necessary to have a base line of discrimination which presumably reflects the discriminability of the speech stimuli "in the raw"—that is, before language learning has taken place. The desire for a base line that, on the one hand, can be obtained from adults, who already have acquired speech discriminations, and, on the other hand, has some face validity as an index of discrimination in its pristine form, places two mutually incompatible requirements on the set of control stimuli used to obtain this base line: First, the control stimuli should be comparable to the speech stimuli, varying in the same acoustic properties

and with the same constant parameters, so that the difference in discriminability of the speech and control stimuli may not be attributed to a difference in the psychophysical task; second, the control stimuli must not evoke the speech discriminations under study, so that their discriminability may indeed serve as a base line for assessing the effects of training in the speech discrimination. In practice, the control stimuli that are employed represent a compromise between these two sets of requirements; usually some of the constant parameters of the speech stimuli are altered so that the transmuted series is not perceived as speech but varies in the same parameter as the speech stimuli.

Evidence for acquired distinctiveness of the speech sounds, which is interpreted as favoring the motor theory of speech perception, could not be obtained in the study of the phonemes /b/-/d/-/g/ described above because a suitable set of control stimuli was not prepared. In experiments with the contrasts /do/-/to/, /ræbɪd/-/ræpɪd/, and /slɪt/-/splɪt/, summarized below, the obtained discrimination functions provided all three kinds of evidence favoring the motor theory of speech perception: increased discriminability at the phoneme boundaries, close correspondence between obtained and predicted discrimination functions, and much greater discriminability of the speech stimuli than their nonspeech controls, showing acquired distinctiveness.

In the experiment employing speech stimuli that ranged perceptually from /do/ to /to/ (Liberman, Harris, Kinney, & Lane, 1961), the stimuli were arranged in one-, two-, and three-step comparisons in ABX triads of the form: ABA, ABB, BAA, and BAB. The discrimination functions obtained from 11 listeners agreed well with

those predicted from their labeling gradients. "Discrimination was better for a given stimulus difference when the two stimuli lay on opposite sides of a phoneme boundary than when the stimuli were within the same phoneme category [p. 384]."

In order to obtain a set of non-speech control stimuli that differed with respect to cues comparable to those defining the speech continuum, the authors inverted their speech spectrograms before converting them to sound with the Pattern Playback.⁶ The control stimuli were prepared in ABX triads according to the same procedure followed for the speech stimuli and were presented a like number of times to the same listeners employed for the labeling and discrimination phases of the experiment. The discrimination functions that were obtained for the control stimuli did not depart appreciably from chance, they did not show any regularly occurring peaks, and they lay well below the speech-discrimination functions in all comparisons.

The authors conclude that, to the extent that the control stimuli provided a valid measure of the discriminability of the speech stimuli prior to language learning, the speech-discrimination functions reflect acquired distinctiveness to a large degree. Two other sources of evidence may be cited in support of this conclusion. Hirsch (1959) had listeners judge the temporal order of two pure tones of differing frequency. The difference in time of onset to give 75% correct judgment was about 17 milliseconds. In comparison, the difference in time of

onset of the first and second formants of the synthetic speech stimuli that gave 75% correct discrimination was about 12 milliseconds. Thus, the peak in the /do/-/to/ discrimination function again may be considered to reflect acquired distinctiveness. Lane and Moore (1962) obtained a "nonspeech" control base line for the /do/-/to/ discrimination by measuring, with an aphasic subject, the discriminability of the same set of *speech* stimuli employed by Liberman, Harris, Kinney, and Lane (1961). Chance levels of discrimination (and a flat labeling gradient) were observed prior to discrimination training. Following training, a gradient of phoneme identification was obtained, and ABX discrimination was well above chance levels in all three-step comparisons, peaking at the phoneme boundary.

In the experiment employing 12 speech stimuli that ranged perceptually from /ræbɪd/ to /ræpɪd/ (Liberman, Harris, Eimas, Lisker, & Bastian, 1961), the stimuli were arranged in two- through nine-step comparisons in ABX triads. The discrimination functions obtained from 12 listeners, shown by the filled circles in Figure 6, reveal a peak at the phoneme boundary (cf. Figure 4) and, in general, agree well with those functions predicted (unfilled circles) from the labeling gradients. The obtained discrimination functions lie, at all points, somewhat above those predicted, suggesting that the listener can discriminate between two speech stimuli slightly more often than he can identify them absolutely.

The speech-discrimination functions shown in Figure 6 are not, however, unimodal. There is one peak in the region of 70 milliseconds corresponding to the phoneme boundary shown in Figure 4 and a second peak in the region of 100 milliseconds that does not correspond to a phoneme boundary.

⁶ Certain additional modifications were made in the control patterns "to provide a further precaution against the possibility that the control stimuli would resemble speech [Liberman, Harris, Kinney, & Lane, 1961, p. 381]."

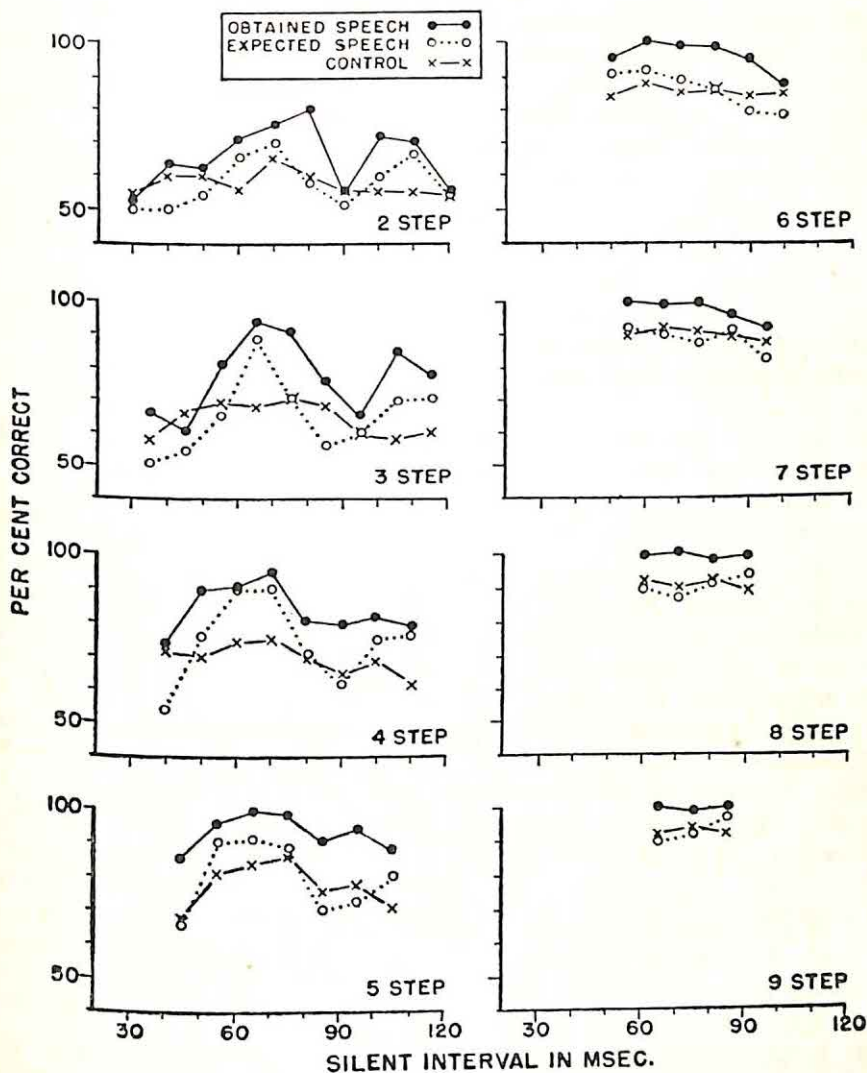


FIG. 6. Discrimination of the synthetic speech stimuli /ræbɪd/-/ræpɪd/ (filled circles) and comparable nonspeech stimuli (crosses), when presented in ABX triads to 12 listeners. (The "expected" functions are based on the labeling gradients obtained from these subjects when they were instructed to identify each stimulus as /b/, /p/, or */p/. Each point on the obtained functions shows the mean relative frequency of correct discrimination in 40 presentations of each triad. From Liberman, Harris, Eimas, Lisker, & Bastian, 1961.)

The authors considered that the stimuli with greater than 100-millisecond silent interval, labeled by their listeners as /p/, were "unnatural." They thought that this phoneme "might nevertheless be heard and articulated by our listeners almost as if it were a different speech entity [p. 186]." The listeners

were then recalled and instructed to label the speech stimuli anew, this time with three possible responses: /b/, /p/, and */p/. The labeling function for the group of subjects then showed a third category, */p/, and hence a second "phoneme boundary": /p/-*/p/. The predicted discrimination

functions shown in Figure 6 were derived from this revised labeling gradient.

In order to assess the degree of acquired distinctiveness shown by the speech-discrimination functions, control stimuli were prepared by processing the speech stimuli so as to yield noise bursts with the same amplitude envelope and the corresponding silent intervals. These control stimuli were arranged and presented for discrimination measurement in the same manner as the speech stimuli. The discrimination of the control stimuli proved inferior to that obtained with the speech stimuli (filled circles) in almost all comparisons, as shown in Figure 6. "The entire learning effect consists of a sharpening of discrimination in the vicinity of the phoneme boundary [p. 189]." This conclusion agrees with that reached earlier in the experiment on discrimination of the phonemes /d/-/t/.

In the experiment employing speech stimuli that ranged perceptually from /slrt/ to /splrt/, reported by Bastian et al. (1961), ABX discrimination was also measured and found to show peaks in the region of the phoneme boundary and, over all, a close correspondence to the discrimination functions predicted from the phoneme-labeling gradients. The acquired distinctiveness of the /slrt/-/splrt/ stimuli was not determined.

The following kinds of evidence, viewed by the experimenters as favoring the motor theory of speech perception, have been cited in this section: When the frequency locus of a schematic stop was held constant and the following vowel varied, consonant identification varied in a way that seemed to accord with the varied patterns of articulation necessary to produce the comparable speech stimuli in natural language. When synthetic speech stimuli were selected from an acoustic continuum that ranged perceptually from one phoneme to another and

these stimuli were presented to listeners for identification, the phoneme-labeling gradients were quite steep, suggesting categorical perception of the speech stimuli, mediated by the discretely different articulations required to produce the two phonemes in natural language. When the synthetic speech stimuli employed for the labeling experiments were arranged in ABX triads for discrimination measurement, listeners more readily discriminated between stimuli drawn from opposite sides of the phoneme boundary than between stimuli drawn from the same phoneme category, suggesting that discretely different mediating responses were involved in the former discriminations, enhancing discriminability by their very different proprioceptive returns. These discrimination functions accorded well with those predicted from the labeling gradients on the assumption that the listener could discriminate speech stimuli only to the degree that he labeled them differently, suggesting that discrimination was controlled, not by the continuously varying speech stimuli, but rather by the proprioceptive return from only two different mediating responses. Finally, when a base line of discriminability was obtained for the speech stimuli by presenting appropriate control stimuli for discrimination measurement, the speech discriminations were found to be much more accurate in all comparisons, suggesting acquired distinctiveness of speech stimuli, brought about by the mediating articulations that these stimuli evoke.

Diverse Acoustic Stimuli, Similar Perceptual and Articulatory Responses

Liberman (1957) has put the question: "When articulation and sound wave go their separate ways, which way does the perception go? [p. 121]." His answer, favoring the motor theory of speech perception, was: "The perception always goes with articulation [p. 121]." Part of the support for this answer is provided by the experiments discussed above, in which diverse articulations produced similar acoustic stimuli that evoked diverse perceptual responses. Additional support for the answer is provided by the complementary findings, discussed below: Similar articulations produce di-

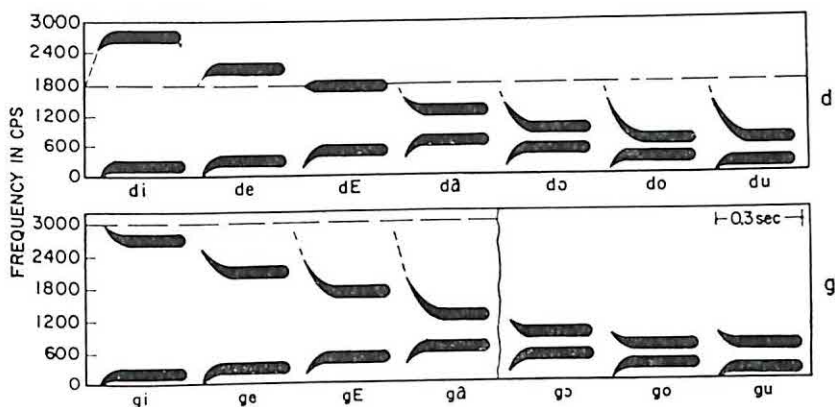


FIG. 7. The spectrographic patterns that produce /d/ and /g/ before various vowels. (The dashed lines are extrapolations to the /d/ and /g/ loci. From Liberman, 1957.)

verse acoustic stimuli that evoke similar perceptual responses.

The first evidence of this sort comes from a study of the perception of the stop consonants discussed earlier (Liberman et al., 1952). When the frequency locus of a schematic stop and the formant frequencies of a following vowel were both allowed to vary, the greatest number of /k/ identifications was not associated with any one schematic stop, but varied widely, depending on the following vowel; for example, a schematic stop at 2,520 cycles per second gave the most /k/ judgments if the following vowel was /e/, but one at 710 cycles per second gave the most /k/ judgments if the following vowel was /u/. The nearest spoken equivalents to the schematic /k/'s are all produced by similar articulatory movements (although the point of the tongue tends to be slightly back for back vowels and more forward for front vowels). "We may say then, in regard to schematic stops heard as *k*, that the perception reflects the basic similarity in the articulatory patterns, rather than the relatively gross differences among the acoustic stimuli [p. 514]."⁷

⁷ "The anomalous status of /k/ is presumably to be explained on the basis of some dis-

Additional evidence seen as favoring the motor theory of speech perception is provided by experiments on the perception of the stop consonant /g/ in initial position. Liberman, Delattre, Cooper, and Gerstman (1954) have shown that the direction and degree of second-formant transition serves as a cue for the perception of the stop consonants, and that the degree or rate of transition necessary for modal identification of a given stop consonant in a CV pair depends on the vowel that follows it, as shown in Figure 7. Therefore, the perception of the consonant remains invariant from one CV pair to the next, despite large changes in the formant transitions. With respect to the stop consonant /d/, the investigators were able to associate the perceptual invariance with an acoustic invariance: Different /d/ transitions before different vowels "seem to be coming from the same frequency position—namely the /d/ locus at 1800 cps [Liberman, 1957, p. 121]." In the

continuity, not at the articulatory level, but at the point of conversion of the articulatory gesture to acoustic output. It would appear that the acoustic description provides a poorer approximation to the ideal phoneme than does the articulatory characterization, at least for /k,g/ [Lisker, Cooper, & Liberman, 1962, p. 94]."

case of /d/, therefore, perception, articulation, and the acoustic stimulus all "go" the same way. In the case of /g/, however, no acoustic invariance was found to be associated with the invariant perception (or with the similar articulations that produce /g/ before the several vowels). Figure 7 shows that this consonant requires a large transition originating at about 3,000 cycles per second when the following vowel is /i/, /e/, /ɛ/, or /a/. When the following vowel is /ɔ/, /o/, or /u/, however, the transition is, on the contrary, very small, and the locus has shifted markedly.

Here, the large and abrupt change in acoustic pattern contrasts with the comparatively small and continuous shifts in articulation. . . . The identity of the perceived consonant remains the same throughout, and thus parallels the invariance of the articulation [Cooper, Liberman, Harris, & Grubb, 1958, p. 937].

Perhaps the most striking invariance in the perception of speech under wide changes in the acoustic stimulus is the constancy of vowel identification despite changes in the vowel-formant frequencies, due to differences among speakers, dialects, and phonetic environments. If each vowel phone is represented as a point on a vowel-formant chart, whose coordinates are the first- and second-formant frequencies of the vowels, and if the set of points corresponding to each phoneme is considered to define a vowel region, then the regions of different vowels sometimes overlap, and the regions of the same vowels sometimes fail to overlap. Even under optimal conditions, when vowels are uttered in neutral phonetic contexts by practiced speakers and only those vowels are analyzed that are identified unanimously and consistently by a group of listeners, even then, the assignment of a vowel phone

to a phoneme class is ambiguous, as shown in Figure 8.

In order to account for the perceptual constancy of the vowel phoneme despite wide variations in its acoustic properties, Joos (1948) put forth, in a classic monograph, a motor theory of vowel perception. Joos notes, first of all, that it is possible to find some order among the vowel regions solely on the basis of their acoustic parameters. This ordering may be described graphically as a vowel triangle⁸ imposed on the formant plot; its apices are the vowel regions /i/, /a/, and /u/. When such a reference system is drawn for the vowels of a particular speaker, most of the other vowels he produces lie, to a degree, on the borders of or within the triangle and in relatively predictable positions. The triangle that is fit to the utterances of a particular speaker, however, differs from one speaker to the next. Joos notes that the between-speakers variance in vowel regions may be greatly reduced by "sliding" the several vowel triangles into congruency, and he describes a statistical procedure for approximating this. Joos suggests that this procedure is a model of the process by which phonemic constancy is achieved in perception. He contends, furthermore, that this normalizing process in vowel perception is controlled, not by the speech signals themselves, but by their articulatory correlates:

The brain, then, contains an equivalent of Fig. 29 [a schematic of the organs of articulation] and upon this map it can lay down an equivalent of Fig. 26 [a vowel quadrilateral with the coordinates: height and position of the tongue-hump during articulation], and if it has two of the latter it can shift and distort one of them to fit the other [p. 63].

⁸ Sometimes a vowel quadrilateral is used as a model. This does not change Joos' (1948) theory substantially.

in terms of their more diverse acoustic parameters.

Changing Perception Correlated Less with the Acoustic Stimuli, More with the Patterns of Articulation that Produce those Stimuli

The four experiments summarized below have in common that the acoustic variable was the sound pressure level of the utterance and the perceptual response was some estimate of loudness or stress.

The suggestion that the perception of a suprasegmental feature of language such as stress may involve mediating articulation was perhaps first made by Jones (1932) who wrote:

Accent is *sui generis* depending for its perception on the kinesthetic sense. . . . The listener refers what he hears to how he would say it. Thus he translates exteroceptor into proprioceptor sensations, the kinesthetic memory serving as stimulus [p. 74].

In a study conducted by Ladefoged, Draper, and Whitteridge (1958), the sentence: "The old man doddered along the road" was uttered by one speaker while the action potentials of his internal intercostal muscles were measured. The authors report that listeners' estimates of the stress loci in the sentence corresponded to the temporal loci of heightened muscular activity. They suggest [p. 9] that statements about stress "are usually best regarded as statements about the speaker's muscular behavior (or about the actions of the listener's muscles which would have to be made in order to produce similar sounds)."

In a subsequent study by Draper, Ladefoged, and Whitteridge (1959), speakers were instructed to count to 20, making the words equally loud. The authors report that the peak intensities of the words differed by as much as 15 decibels, "but subjects were not usually aware of these differences

[p. 21]." Although a speaker's vocal level varied, the mean level of air pressure in his lungs, found in other measurements to depend on "the loudness with which the speaker was trying to talk," remained "fairly constant" under instructions to count to 20 with equal loudness. The authors conclude that "naive listeners, obeying an instruction to consider the loudness of sounds in continuous speech, do not assess the acoustic properties of the sounds but consider, instead, the pressure which would be required below the vocal cords [p. 21]."

The most carefully controlled experiment whose findings are seen as favoring the inference of mediating articulation in the perception of stress was performed by Ladefoged and McKinney (1963). The five words: BEE, BAR, BAY, BORE, and BOO were uttered by one speaker at various vocal levels under conditions that permitted measurement of the peak sound pressure and peak subglottal air pressure during the voicing of each utterance. These utterances were tape-recorded, and each was spliced into the recorded carrier phrase: "Compare the words BAR and. . . ." Thirty listeners were presented with 65 sentences constructed in this way. The subjects were instructed to consider the (constant) loudness of the word BAR in the carrier to have a magnitude of "10" and to state the relative loudness of the terminal word by assigning it a number in the same proportion to 10 as its loudness was to the standard, BAR. The means of the loudness estimates (L) were found to be related to the sound pressure level (S) by a power function with exponent 1.2 ($L \propto S^{1.2}$). The loudness estimates of the words may also be related to the work (W) done upon the air in producing the voiced sounds. This work is proportional to the product of the subglottal

pressure and the rate of flow of air through the glottis ($W \propto P \cdot F$). In an earlier experiment of this series, the authors showed that $P \propto F$ (hence $W \propto P^2$) and that $P \propto S^0$. The last two expressions give $W \propto S^{1.2}$. Since $L \propto S^{1.2}$, the loudness estimates of the speech stimuli were linearly related to the work done in producing them. The authors conclude: "It would appear that the perceived loudness of words which are within the normal speech range is largely dependent on . . . the physiological effort required to produce them [p. 459]."

Lane, Catania, and Stevens (1961) have reported that the speaker's estimates of his own vocal level grew as the 1.1 power of the sound pressure he produced. As Ladefoged and McKinney (1963) point out, their finding that $W \propto S^{1.2}$ suggests that the Lane et al. (1961) speakers were judging physiological effort. This hypothesis is supported by the finding that the "autophonic scale," $\psi \propto S^{1.1}$, was not changed when the speaker was deprived of auditory feedback by intense masking noise. Ladefoged and McKinney (1963) conclude that the loudness of a speech sound, whether generated by the listener himself or by an external source, is perceived in terms of the physiological effort required to produce that sound.

The last study of the perception of speech loudness to be cited in support of a motor theory of speech perception was reported by Lehiste and Peterson (1959). A tape recording was prepared in which various vowels produced with equal effort were mixed at random with vowels produced with unequal effort. The vowels were arranged in pairs at random and presented at the *same* sound pressure level to listeners instructed to judge the louder of the two vowels in the pair. "Almost invariably, the listeners iden-

tified the vowels that were produced with a greater amount of effort as louder than vowels having greater intrinsic amplitude but produced with normal effort [p. 431]." In the light of these findings, the authors suggest that:

The listener interprets speech according to the properties of the speech production mechanism rather than according to the psychophysical principles of the perception of abstract sounds [p. 428].

The following kinds of evidence, viewed by the experimenters as favoring the motor theory of speech perception, have been cited in this section: When listeners judge the stress loci in a sentence, these judgments correlate most closely with temporal loci of heightened muscular activity in the speaker, suggesting that the listener is guided by covert articulatory cues in making his estimates. When speakers are instructed to count, making the words equally loud, large changes in intensity are produced although the mean level of air pressure in the lungs remains constant, suggesting that the speaker's criteria for equal loudness depends on proprioceptive cues from muscles involved in articulation. When listeners estimated the loudness of a word uttered at various levels, their estimates were linearly related to the work done in uttering the variable word; these estimates were also related linearly to the autophonic scale which describes the speaker's perception of his own vocal level. Inference from these findings suggests that the listener judges the loudness of speech sounds in terms of the effort cues associated with their production. Finally, when pairs of vowels uttered with unequal effort were presented at equal sound pressure levels and listeners judged the louder vowel in each pair, the vowels produced with greater effort were always judged louder, suggesting that the perception of loudness is more closely correlated with changes in articulation than with the acoustic stimulus.

EVIDENCE THAT OPPOSES THE MOTOR THEORY OF SPEECH PERCEPTION

The inference that speech is perceived by reference to articulation was arrived at, in the studies discussed previously, after examining principally three dependent variables: identifica-

tion, discrimination, and loudness-estimation functions. The properties of these functions that were repeatedly cited by the investigators as favoring a motor theory of speech perception were: the steepness of the identification gradients, the correlated peaks in the discrimination functions, the accuracy with which the discrimination functions may be predicted from the labeling gradients, and the correlation of loudness estimates with vocal-effort measurements.

This section of the review considers some methodological limitations of the studies cited above and presents evidence that the reported relations between identification and discrimination functions are not at all unique to the perception of consonants but describe as well the perception of vowels and of entirely nonlinguistic stimuli such as complex tones, sectors of circles, and patches of color. Clearly, these findings with nonlinguistic stimuli militate against the postulation of a special perceptual mechanism for speech. The same findings for the vowel stimuli also contravene the motor theory for this reason: The theory maintains that the perception of most consonants is discontinuous or categorical because they covertly evoke discontinuous articulations whose discretely different sensory feedback controls the overt identifications, whereas the perception of vowels is continuous because the mediating articulations are continuous and provide similar sensory returns.

Finally, this section shows that the sensory dynamics of listening and speaking are drastically different, disputing the purportedly close relation between the perception of loudness and of vocal effort.

Identification Functions for Consonants, Vowels, and Nonspeech Stimuli

The following considerations preclude construing identification functions as providing support for a motor theory of speech perception:

1. The widespread practice of preselecting subjects and of presenting their pooled data causes the form of the identification functions associated with a stimulus continuum to vary arbitrarily.

2. It is not possible in principle to compare the steepness of vowel and consonant labeling gradients in some nonarbitrary way since their abscissae are noncommensurate.

3. A direct comparison of the labeling gradient for a consonant continuum with that for a commensurable nonlinguistic continuum reveals that the gradients are comparably steep.

Subjects were selected for participation in several of the experiments reported above depending on the steepness of their labeling gradients, as measured on a pretest. For example, in the experiment employing the phoneme contrast /do/-/to/, "the 11 Ss [from a total of 20] whose judgments provided the sharpest phoneme boundaries were selected for the experiment [Liberman, Delattre, & Cooper, 1958, p. 382]"; in the /b/-/d/-/g/ experiment, "two of the five Ss in group I were eliminated because they failed to apply phoneme labels consistently [Liberman et al., 1957, p. 360]"; and in the /b/-/p/ experiment, "there were 12 subjects in all, chosen from a group of 31 on the basis of a pretest. The purpose of the selection was to insure that all subjects ultimately serving in the experiment would have a *sharp and clear* phoneme boundary [Liberman, Harris, Eimas, Lisker, & Bastian, 1961, p. 181; *italics mine*]."

The rejection of 50% of the subject population in these three studies according to a criterion that was neither explicit, quantitative, nor uniform greatly restricts the generality of inferences based on the findings obtained with the remaining subjects. This procedure of preselecting subjects vitiates, moreover, comparisons among different experiments with respect to the form of the obtained labeling gradients, the predicted discrimination functions which are derived from these gradients, and, therefore, the disparity between predicted and obtained discrimination functions. As a result, it has not been possible to assess the accuracy of an important prediction that follows from the motor theory of speech perception; namely, that perception of the consonants is more categorical than is perception of the vowels.

The procedure of averaging individual identification and discrimination functions and presenting group data, employed in all the synthetic speech studies cited earlier but two, may be placed alongside preselection of subjects in vitiating comparisons between the findings of experiments when these comparisons are based directly or indirectly (see later) on the labeling gradients. Since it is generally true that the boundary of the labeling gradient varies from one subject to the next in its locus along the stimulus continuum, the average gradient is considerably less steep than the individual gradients—and reflects none of them well. Put in other terms, this procedure confounds the within-subjects variance, which relates to the judgment of categorical perception with the between-subjects variance, which does not.

Comparisons of identification gradients encounter difficulty even when individual gradients for randomly selected normal adults are considered. When two gradients, both functions of

the *same* synthetic speech continuum, are compared, the judgment of categorical perception may be based on the extent of the plateaus in response probability or, equivalently, the steepness of the boundary between these plateaus. However, the steepness of a labeling gradient or, derivatively, the height of the predicted discrimination function for the same stimulus continuum, is of interest largely in relation to the steepness of other gradients collected with *other* stimulus continua: If the steepness of consonant-labeling gradients or the height of discrimination functions derived from them is cited as evidence of discretely different mediating responses, then it must be shown that these gradients are more steep than those obtained in comparable experiments with vowels, whose articulations are topographically continuous, and certainly more steep than those obtained with nonspeech stimuli.

When comparing labeling gradients for *different* stimulus continua, one gradient may not be called more categorical than the other solely because it has more extended regions over which response probability is zero or unity; the extent of these plateaus in the labeling gradient is largely a matter of experimental convenience. For example, in the experiment whose outcome is shown in Figure 4, the silent interval for the rapid-rapid continuum could have been allowed to vary only over the range 40 to 90 milliseconds. The outer boundaries of the phoneme classes have not been defined experimentally, so it is not possible to say to what extent the gradients "partition" the total perceptual range.

It appears that the sole property of diverse gradients that will reliably serve as a basis for judgments of steepness is the value of their slope over the range in which it is not zero. However, the slopes of the diverse gradients

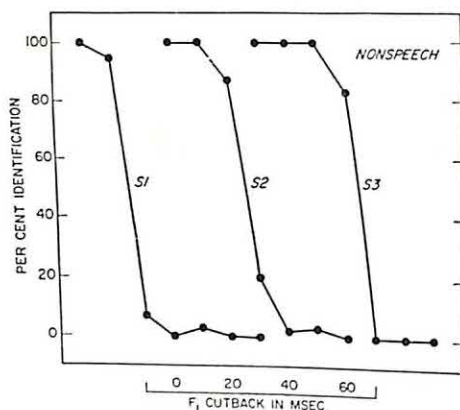


FIG. 9. Identification of nonspeech stimuli. (The seven spectrographic patterns in the series employed by Liberman, Harris, Kinney, & Lane, 1961, to study /do-/to/ identification were inverted and then converted to sound. Each of three subjects was conditioned to respond to the stimulus with no cutback of the first formant and to the stimulus with 60-microseconds cutback. In generalization testing, each of the seven stimuli was presented 60 times, in irregular order.)

(or the entire functions for that matter) are obviously not comparable—despite the widespread practice of comparing them—since their abscissae are not commensurate.⁹ It may still be possible to compare diverse gradients if some normalizing procedure is found that is not sensitive to arbitrary differences between the experiments, such as the number and spacing of the stimuli employed. That there exists a satisfactory procedure may be strongly doubted. In any event, direct comparisons of functions with noncommensurate abscissae are out of the question albeit in vogue.

There is considerable evidence to support the conclusion that the general

⁹ As an explicit example of these improper comparisons: "Identification of the consonants was marked by abrupt transitions . . . Vowel perception was, however, continuous; the identification functions showed less steep boundaries . . . [Eimas, 1963, p. 206]."

form of labeling functions for consonant continua does not differ from that for vowel and nonlinguistic continua. Clearly defined plateaus and boundaries have been observed in vowel identification by Studdert-Kennedy, Liberman, and Stevens (1963); Stevens, Ohman, and Liberman (1963); Bastian and Abramson (1962); and Fry, Abramson, Eimas, and Liberman (1962), among others, while these properties have been found with identification functions for nonlinguistic continua by Cross, Lane, and Sheppard (1965); Lane and Kopp (1964); Beare (1963); and Cross and Lane (1962), among others.

There is also evidence that the detailed form of labeling gradients for consonant and nonlinguistic continua do not differ appreciably. Such comparisons are appropriate only if the stimulus variable is the same in the two cases. Therefore, Lane and Schneider (1963) selected the published labeling gradients for /do-/to/ as representative gradients for consonant continua and then employed the /do-/to/ *control stimuli* to obtain comparable gradients for a nonlinguistic continuum. It will be recalled from the earlier discussion of the /do-/to/ study by Liberman, Harris, Kinney, and Lane (1961) that the control stimuli were obtained by inverting the /do-/to/ spectrograms before converting them to sound (an operation which preserves the stimulus variable), that they were not perceived as speech stimuli, and that their discrimination functions showed only chance levels of accuracy. The two extreme stimuli from this series (with 0 and 60 milliseconds of cutback in F1, respectively) were presented in random order to each of eight listeners serving individually. The subject was seated in front of a tape recorder and counter and was visually isolated from the experimenter.

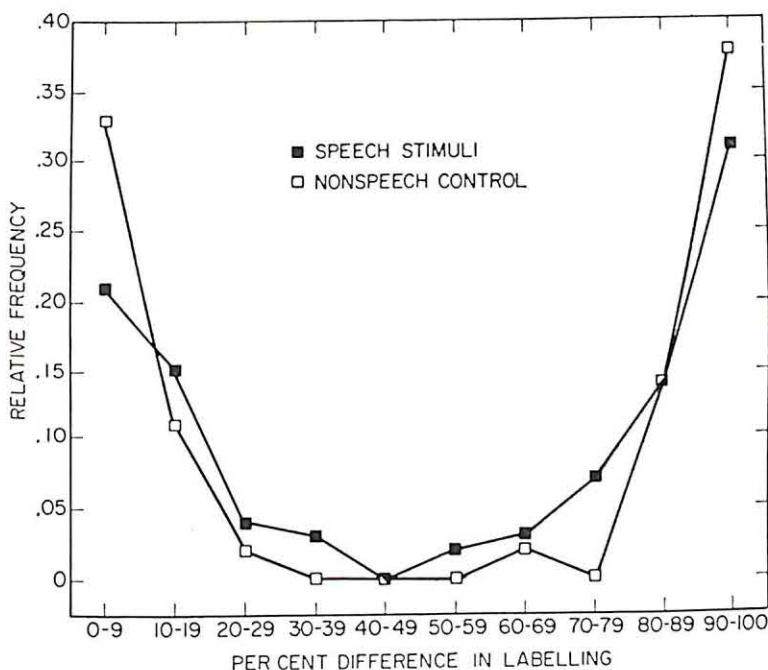


FIG. 10. Comparison of the degree of "categorical perception" obtained in identification of synthetic speech stimuli (Liberman, Harris, Kinney, & Lane, 1961) and comparable nonspeech stimuli (Lane & Schneider, 1963). (For each experiment, each stimulus was compared to every other in percent identification, and the relative frequency of within-pairs differences was plotted. In the limiting case of "categorical perception," the labeling gradients would have the form of a step function, and the frequency distribution would be nonzero only in the extreme classes.)

He was told to say /do/ or /to/ in each 4-second interstimulus interval, that he would receive a point on the counter for each correct identification, and that each point was worth a penny at the end of the experiment. The experimenter, guided by a protocol, presented the point immediately after each correct response. Conditioning was terminated after either 30 correct identifications in succession or 400 stimulus presentations. Three of the eight subjects who participated met the former criterion and were presented individually with the generalization-test series, in which the seven control stimuli occurred in an irregular order, with the constraint that each stimulus followed every other exactly twice. In the second through fourth daily experi-

mental sessions, conditioning was reduced to 100 stimulus presentations, and generalization testing was replicated.

The generalization gradient for the nonspeech (control) stimuli, obtained from each of the three subjects, is shown in Figure 9. These gradients are directly comparable to the /do/-/to/ labeling gradients reported for 13 subjects by Liberman, Harris, Kinney, and Lane (1961). This comparison reveals that the steepness of nonspeech gradients may equal or exceed that for speech gradients. One way of combining the individual gradients in each of these two studies that is not sensitive to differences in the loci of the gradients on the stimulus continuum is shown in Figure 10. To illustrate how

the functions shown there were obtained, the procedure employed with the individual gradients in Figure 9 will be described. For each subject, the relative frequency with which the /do/ response was emitted to each of the seven nonspeech stimuli was noted as a percentage. Each stimulus was compared to every other in the percent of /do/ responses it evoked, and the difference in percent was obtained for each pair. A frequency distribution of these percent differences was prepared, with class intervals of 10%, and the frequency distributions for the three subjects were combined. The frequency measures were then converted to relative frequency and plotted in Figure 10. If the individual gradients were step functions, the limiting case of "categorical perception," then all the percent differences would be either 0 or 100, and the frequency distributions would have peaks in the extreme classes leaving intermediate classes empty.

Comparison of the frequency distributions in Figure 10 for identification of synthetic speech stimuli (Liberman, Harris, Kinney, & Lane, 1961) and groups of tones differing in relative onset times (Lane & Schneider, 1963) yields no evidence for the claim that the perception of speech stimuli is peculiarly categorical.

Discrimination Functions for Consonants, Vowels, and Nonspeech Stimuli

The relations between discrimination functions obtained for synthetic speech continua, those predicted from the corresponding identification gradients, and those obtained with comparable nonspeech control stimuli force an extra degree of complexity on the motor theory of speech perception, which must introduce a direct route for speech discriminations in order to account for the relations among these functions.

Concerning the first relationship, illustrated in Figure 6, the obtained discrimination functions lie above those predicted from the identification gradients. This finding has been interpreted as showing that "listeners are able in discriminating these stimuli to extract some information in addition to that which is revealed by the way in which they label the stimuli as phonemes [Liberman, Harris, Eimas, Lisker, & Bastian, 1961, p. 185]." This interpretation implies that the listener's speech discriminations are controlled by mediating articulations *and* by the acoustic stimulus, either alternately or concurrently, and in some relative degree.

Concerning the second relationship, also illustrated in Figure 6, the speech-discrimination functions lie above those for the nonspeech control stimuli, which are purportedly comparable, both at the phoneme boundary and within phoneme categories. If the listener were relying exclusively on articulatory cues in making the speech discriminations and if the articulations were truly categorical (i.e., identical within phoneme classes), then the speech-discrimination functions should fall, within phoneme classes, at the level of or below those functions for the control stimuli. That they do not again forces the interpretation—if we are to preserve the mediation hypothesis—that the listener sometimes utilizes proprioceptive information and sometimes does not. It is unclear by what means the listener manages to use this information just when it has the most to offer, at the phoneme boundary, and to ignore this information when it would detract from discrimination, within phoneme classes.

A consideration of the relations among discrimination functions for consonants, vowels, and nonspeech stimuli provides further evidence that

milittates against the motor theory of speech perception. Concerning the first comparison, it has been found that, although consonants are generated in natural language by discontinuous articulations and vowels are generated by continuous articulations, enhancement of discrimination at the phoneme boundary occurs with both consonant and vowel stimuli. This finding contradicts the outcome predicted by the motor theory of speech perception:

When significant acoustic cues that occupy different positions along a single continuum are produced by essentially discontinuous articulations (as, for example, in the case of second-formant transitions produced for /b/ by a movement of the lips and for /d/ by a movement of the tongue), the perception becomes discontinuous (i.e., categorical), and discrimination peaks develop at the phoneme boundary. When, on the other hand, acoustic cues are produced by movements that vary continuously from one articulatory position to another (as, for example, the frequency positions of first and second formants produced by various vowel articulations) perception tends to change continuously and there are no peaks at the phoneme boundaries [Liberman, Harris, Eimas, Lisker, & Bastian, 1961, p. 177].

Employing synthetic vowel stimuli, Liberman, Stevens, and Ohman (1963) have found, on an identification test, "reasonably sharp and well-defined boundaries between phoneme categories," and, on a discrimination test, "peaks at the phoneme boundaries, i.e., discrimination between adjacent stimuli in the vowel continuum is better when the stimuli lie near phoneme boundaries [p. 13]." In the study of synthetic vowels by Fry et al. (1962), a peak in vowel discrimination is predicted in the region of the / ϵ -/- α / boundary for the two- and three-step comparisons. Unfortunately, this peak could not be verified by the obtained discrimination data since discrimination over the entire continuum was nearly 100% correct in these comparisons.

There is, however, some indication of a peak in discrimination in the region of the boundary for the one-step comparisons. To raise the ceiling on vowel discrimination in the latter study, Cross and Lane (1964) copied the tape recordings employed by Fry and co-workers, kindly provided by the Haskins Laboratories, while inserting a broad middle-formant intermediate to the two formant frequencies originally comprising each vowel. The series of degraded vowels was then presented to listeners for labeling and ABX discrimination. With the overall level of discrimination somewhat reduced, marked peaks in the discrimination functions were observed at the phoneme boundaries.

It should be noted, however, that Bastian and Abramson (1962) did not find peaks in the discrimination of Thai vowels, minimally distinguished by duration. Whether this finding, which contrasts with those just reported and with the data for nonspeech stimuli that follows, should be attributed to some unique property of duration as a discriminative variable, or to a procedural artifact such as near perfect discrimination over the entire continuum, cannot be determined from the abstract of the paper they presented.

Concerning the second comparison, cited above, between consonants and nonspeech stimuli, it has been found that the latter also yield enhancement of discrimination at the boundaries of their identification gradients. To return first to the experiment on discrimination of / $ræbrd$ /-/ $ræpɪd$ / (Liberman, Harris, Eimas, Lisker, & Bastian, 1961) described earlier, it will be remembered that a second peak was observed in the discrimination function, and that this peak was subsequently shown to correspond to a third "phoneme category," labeled */p/ ("the unnatural /p/"). The obtained peak

TABLE 1

AVERAGE PERCENT CORRECT DISCRIMINATION
OF NONSPEECH STIMULI, PRESENTED IN
ABX TRIADS TO THREE SUBJECTS

Comparison	For A and B stimuli drawn from the same "phoneme class"*		For A and B stimuli drawn from different "phoneme classes"*		$\mu = d/\hat{\sigma}^{**}$	P
	%	N	%	N		
One-step	49.8	239	65.6	63	2.80	.01
Two-step	66.7	192	80.7	155	3.40	.001
Three-step	71.1	91	85.5	219	3.40	.001

* The difference in percent identification of A and B stimuli was always either less than 25% (same class) or greater than 60% (different classes).

** Normal test of differences between percents, Wadsworth and Bryan (1960).

in discrimination at the /p/-*/p/ boundary is considered, according to the motor theory of speech perception, to reflect differentiated proprioceptive feedback from diverse articulatory patterns evoked by /p/ and */p/. It is difficult to imagine, however, what pattern of articulation has become attached to */p/ over long years of language conditioning, especially since the authors report that "this was a strange and unnatural /p/ to American ears [p. 186]." In this case, too, the hypothesis that speech is perceived by reference to its articulatory correlates does not provide a convincing account of the observed relation between labeling and discrimination.

Three experiments employing exclusively nonlinguistic stimuli provide the most convincing evidence that the observed relation between labeling and discrimination functions is not peculiar to the perception of speech and is not predicated on "long experience with a language [during which] a speaker (and listener) learns to connect speech sounds with their appropriate articulations [Liberman, Harris, Eimas, Lisker, & Bastian, 1961, p. 177]." Lane and Schneider (1963) conditioned two identification responses to two non-

speech auditory patterns (the /do/-/to/ control stimuli) and then measured generalization at intermediate stimulus values according to the procedure described earlier. Prior to identification training in the first experimental session, and after that training in all four sessions, each subject was presented the series of control stimuli arranged in triads for ABX discrimination. Before identification responses were conditioned to the extreme stimuli in the series, the discrimination functions for the nonspeech stimuli showed chance levels of accuracy. Following training, these (ABX) discrimination functions showed peaks at the phoneme boundary in the one-, two-, and three-step comparisons for all subjects. Table 1

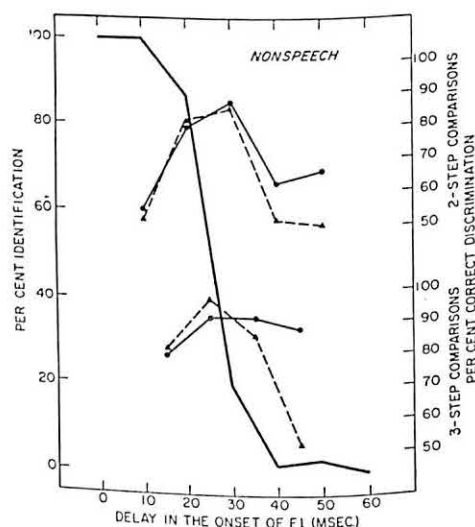


FIG. 11. Discrimination of nonspeech (control) stimuli when presented in ABX triads to one listener. (Each point is the relative frequency of correct discriminations in 24 presentations of pairs of stimuli separated by two or three steps in the series of seven stimuli. This subject's identification gradient, which appears in Figure 9, S2, is also plotted to reveal that peaks in discrimination occur at the "phoneme boundary" of the nonspeech stimuli. The discrimination functions predicted from the identification gradient are shown by the dotted lines.)

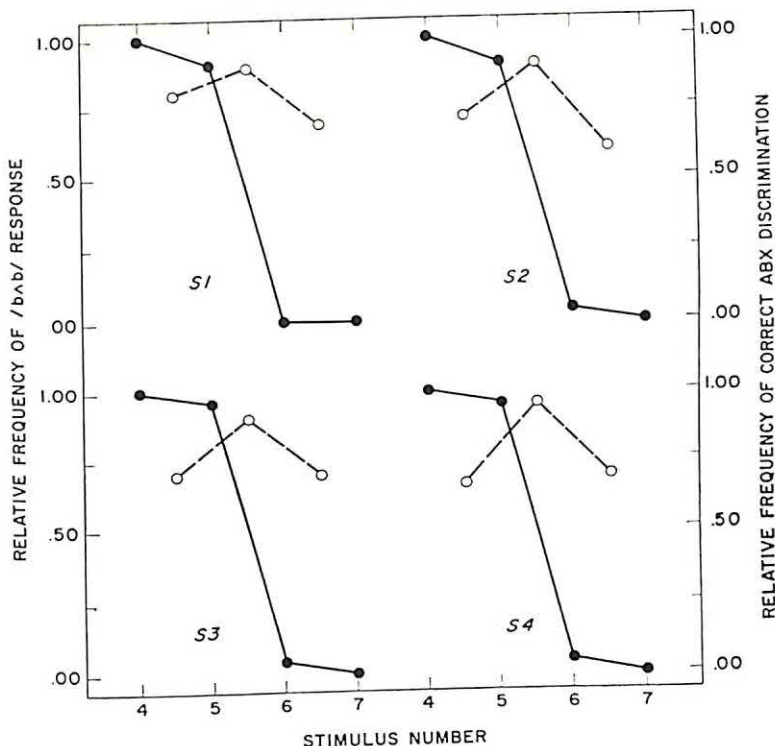


FIG. 12. Identification and discrimination of visual stimuli by individual subjects. (Each point on the identification functions, filled circles, solid lines, shows the relative frequency of the /b^b/ response in 24 presentations of the corresponding stimulus. Each point on the discrimination functions, open circles, dashed lines, shows the relative frequency of correct discriminations in 36 presentations of each pair of stimuli.)

shows that the three listeners discriminated more readily between stimuli drawn from opposite sides of their (nonspeech) "phoneme boundary," shown in Figure 9, than between stimuli drawn from within a "phoneme category." The labeling and ABX functions (solid lines) for the subject who provided the most orderly discrimination data are shown in Figure 11.

Cross et al. (in press) also obtained a correlation between discrimination and labeling employing nonspeech stimuli—in this case visual patterns. These investigators conditioned each of four subjects to respond /b^b/ when circles with deleted sectors of 42° or 46° were presented and /g^g/ when those with sectors of 50° and 54°

were presented. When this labeling repertoire was thoroughly established (50 consecutive correct identifications), the stimuli were presented in a permuted order for identification and then again in triads for ABX discrimination. The outcome is shown in Figure 12: For all subjects the stimuli that evoked different labeling responses were more accurately discriminated than those that evoked the same labeling response.

The correlation between labeling and discrimination has been observed with yet another nonlinguistic continuum: wavelength. Beare (1963) reports labeling gradients for spectral colors that intersect at 450, 490, 575, and 620 mμ. Ekman (1963) obtained functions of similar form when he extracted five factors from a matrix of similarity rat-

ings obtained for spectral colors and then plotted factor loadings as a function of wavelength. The findings of both Beare (1963) and Ekman (1963) show intersections of adjacent gradients at wavelengths that turn out to correspond to minima in the functions relating the size of the jnd to wavelength (Jones, 1917). Thus, the discrimination of wavelength is most acute at the boundaries between color classes.

Since the correlation between discrimination and labeling is observed for nonspeech as well as for speech stimuli, the correlation per se does not seem to warrant the postulation of a special perceptual mechanism for the discrimination of speech stimuli.

Correspondence of Identification and Discrimination Obtained with Consonants, Vowels, and Nonspeech Stimuli

Nor does the degree of correspondence between labeling and discrimination provide a basis for distinguishing among the perception of consonants, vowels, and nonspeech stimuli and therefore a basis for postulating a special mechanism in the case of speech.

It must be acknowledged that an initial comparison of the obtained vowel-discrimination functions and those predicted from the labeling gradients shows a poorer correspondence in the case of vowel stimuli (in the one set of findings published in detail) than in the case of consonant stimuli, as Fry et al. (1962) have noted. However, the apparent difference between the degree to which labeling and discrimination correspond in vowels versus consonants may be attributed to several concurrent artifacts:

1. The predicted discrimination function is derived from the averaged labeling gradients of preselected subjects. As discussed earlier, the criterion for preselection, based on "steepness," has

not been (and probably could not be) comparable in experiments with vowels and consonants. Therefore, the form of the average labeling gradients and the predicted discrimination functions are also not comparable between experiments. Had Fry et al. (1962) rejected not 9 but 10, say, of their 17 subjects, their predicted discrimination functions for the vowels would have been higher; had Liberman, Harris, Eimas, Lisker, and Bastian (1961) retained not 12 but 13 of their 31 subjects, their predicted functions for the consonants would have been lower. Since the experiments are not comparable with respect to the heights of the predicted discrimination functions, they are not comparable with respect to the correspondence of these functions to those obtained.

2. The predicted discrimination function is based on three assumptions discussed earlier, and each of these assumptions is violated by the experimental conditions. The first—that speech stimuli are discriminated to the extent that they are labeled differently—has been disconfirmed by every experiment. This is evidence of what was obvious beforehand, that acoustic continua differ in their psychophysical properties. Contrasting vowels and consonants psychophysically, we note, for example, that the relevant cue may last more than $\frac{1}{3}$ second in the former case, less than $\frac{1}{30}$ second in the latter. Moreover, vowels differ in their "intrinsic" amplitudes (Lehiste & Peterson, 1959) by as much as 6 decibels whereas the stop consonants do not. Fry and his collaborators chose to allow these amplitude fluctuations in their recordings of triads of synthetic speech stimuli, with the consequence that—according to measurements of these tapes in our laboratory—45% of the triads could be discriminated cor-

rectly on the basis of amplitude cues alone.¹⁰ These and other psychophysical cues tend to separate the obtained discrimination functions from those predicted in the case of the vowels but less so in the case of the consonants. Thus, the correspondence between labeling and discrimination may differ among these experiments for reasons that are irrelevant to a motor theory of speech perception.

3. The second assumption on which the predicted discrimination functions are based is also disconfirmed by experimental evidence in the case of vowel discrimination. This assumption provides that the phoneme classes to which stimuli are assigned during the labeling task are the same as those to which they are assigned during the discrimination task. We presented the tape-recorded series of synthetic vowels used by Fry et al. (1962) to a subject in our laboratory with the instructions to identify each stimulus either by circling one of the following words which appeared in a row on his answer sheet, HEED-HID-HEAD-HAD-HOD, or by circling the dash between adjacent words if the sound seemed to fall in a class between them. The result: nearly equal peaks in labeling probability for the last four phoneme classes, with a series of secondary peaks for HEED and the intermediate classes. When the subject is

not seriously constrained in his labeling responses, the number of labeling gradients is largely a function, of course, of the stimulus range, but also, secondarily, of the naturalness of the synthetic stimuli which, in the experiment by Fry et al. (1962), was admittedly "adequate but no more [p. 178]." Had these investigators permitted their subjects a greater number of phoneme labels and then derived the predicted discrimination function accordingly, the disparity between the latter and the obtained function would be considerably reduced (the predicted functions would be elevated). The same may not be said for the consonant contrasts that have been investigated, however. Therefore, the artifact which we have been discussing makes the results for consonants and vowels seem disparate.

4. Finally, the predicted discrimination functions for the vowels violate a third assumption of the prediction formula, "the various stimuli within each triad are perceived independently of each other [Liberman et al., 1957, p. 363]." In the same article in which Fry and others remark on the disparity between obtained discrimination functions and those predicted on the assumption of independence, they present conclusive evidence of the nonindependence of vowel identifications within each triad. Based on a related study by Cross and Lane (1964), we may estimate that had Fry et al. (1962) acknowledged context effects in their predictive formula, the apparent disparity between predicted and obtained discrimination functions would have been reduced 10 to 15%. The finding of greater context effects in the perception of vowels than in the perception of stop consonants is indeed significant, if its generality can be shown, but its relevance for a motor theory of speech perception remains to be stated and

¹⁰ In fact, 82% of the triads showed intensity changes correlated with those in vowel quality. Since the difference limen for vowel intensity is approximately 1.5 decibels (Flanagan, 1955), we counted only those cases in which the correlated intensity change in a triad exceeded 1.5 decibels, yielding a conservative estimate of the number of triads in which intensity cues might have enhanced the obtained discriminations. More recent research on the discrimination of synthetic vowels at the Speech Transmission Laboratory, Stockholm, which also allowed these amplitude fluctuations, found comparable discrimination functions (K. N. Stevens, personal communication, January 1965).

demonstrated. Since the relational nature of vowel perception lies behind the apparent disparity between predicted and obtained discrimination functions and therefore behind the tendency to view vowels as perceived continuously rather than categorically, Lisker et al. (1962) had it just backward when they wrote: "We suspect that the continuous (as opposed to categorical) nature of vowel perception lies behind these examples of the tendency to perceive vowels relationally [Footnote 33]."

Many of these obstacles, which block a direct comparison between consonant and vowel continua with respect to the correspondence of labeling and discrimination, can be circumvented in a comparison of consonant continua with appropriate nonspeech continua. As discussed earlier, Lane and Schneider (1963) carried out such a comparison between the perception of the /do/-/to/ continuum and that of a series of complex tones varying in the same parameter. (The latter were actually the /do/-/to/ control stimuli, speech spectrograms inverted before conversion to sound.) It has been shown earlier in this review that these speech and nonspeech continua do not lead to different labeling gradients or to different discrimination functions. It follows that they do not lead to different degrees of disparity between predicted and obtained discrimination functions. This is illustrated in Figure 11 for the nonspeech case, where it will be seen that the predicted discrimination functions (dotted lines) correspond well to those obtained. Comparison with similar data for the /do/-/to/ continuum reported by Liberman, Harris, Kinney, and Lane (1961) and for the /b/-/d/-/g/ continuum shown in Figure 5 of this review confirms that speech and nonspeech stimuli do not differ in this regard.

Since the degree of correspondence between discrimination and labeling does not differ appreciably for speech and nonspeech continua, the degree of correspondence does not provide a basis for the postulation of a special perceptual mechanism for the discrimination of speech stimuli.

Gradients of Identification Latency for Consonants, Vowels, and Nonspeech Stimuli

Because of the discrimination peaks at the phoneme boundaries, the incoming [consonant] sounds are heard categorically . . . and they are therefore quickly and accurately sorted into the appropriate phoneme bins [Liberman, Cooper, Harris, & MacNeilage, 1963, p. 3].

The correspondence among identification latency, discrimination enhancement, and identification probability, often adduced in characterizing consonant perception,¹¹ describe equally well the perception of vowels and of nonspeech stimuli. This further uniformity in the behavioral effects of the three classes of acoustic continua adds another dimension in which they are congruent.

Cross and Lane (1962) reported a systematic relationship between the relative frequency and the latency of identification responses evoked by stimuli from a synthetic speech continuum, /do/-/to/: Latencies were minimal at stimulus values corresponding to plateaus in identification (the "phoneme classes"), and they were maximal at stimulus values corresponding to the boundary of the identification gradient. Studdert-Kennedy et al. (1963) have reported comparable findings, both for synthetic speech consonants, /b/, /d/, /g/, and vowels, /i/, /I/, /ε/: "For both stops and vowels, voice and

¹¹ Liberman, Cooper, Harris, and MacNeilage (1963); Liberman, Harris, Eimas, Lisker, and Bastian (1961); Eimas (1963).

button-press reaction times were lower near phoneme centers than in the vicinity of phoneme boundaries [p. 33]."

Gradients of identification latency for consonant and vowel stimuli are comparable not only to each other but also to latency functions for nonspeech stimuli. Cross and Lane (1962) reported gradients of identification latency for pure-tone intensities that also reached minima within identification plateaus and maxima at the boundary of the identification functions. Lane and Kopp (1964) later replicated these findings. Beare (1963) obtained comparable findings for the relative frequency and latency of responses during color naming. A psychophysical study by Kellogg (1931) obtained a similar relation between identification frequency and latency. Kellogg used seven fixed pairs of visual intensities that varied in the luminance ratio between the members of each pair. His subjects responded either "left side darker" or "right side darker" to each pair, and latencies were recorded. When the probabilities of the two responses approached equality, latencies were maximal; when response probabilities were most disparate, latencies were minimal.

In the experiment on identification of visual forms cited earlier, Cross et al. (1965) also determined the latencies of their subjects' responses which constituted either a consonant contrast, /b \wedge b/-/g \wedge g/, or a vowel contrast, /b \wedge b/-/bib/. These investigators conclude:

The response-latency functions obtained in this study, with both vowel-contrast and consonant-contrast labeling responses, reach a maximum at the "boundary" of their corresponding identification gradients and a minimum within the plateaus of those functions. . . . These findings for response latency in identification of nonspeech visual forms accord with those reported for identification of synthetic speech stimuli . . . [p. 72]

The same relationship between identification frequency and latency for nonspeech stimuli was obtained by Cross and Lane (1962) under conditions that permitted measurement of changes in response topography. Two vocal responses whose articulations were continuous were employed: humming with a fundamental frequency of 147 cycles per second and humming with a fundamental frequency of 227 cycles per second. These responses (± 2 cycles per second) were first differentiated out of the subjects' vocal repertoires by selective reinforcement, and then conditioned discriminatively to noise bursts at 50 and 80 decibels (SPL), respectively. Measurements of response probability, latency, and topography (fundamental frequency), made during subsequent generalization testing, are represented for three subjects in Figure 13. Inspection of the figure reveals, first of all, the characteristic relations between identification frequency and latency obtained in studies employing consonant, vowel, and nonspeech continua. Most important, it will be seen that there was no tendency for the responses to change topographically with changes in the stimulus. This finding with nonspeech stimuli accords with the outcome of an experiment, discussed earlier in this article, by Harris et al. (1961). These investigators made electromyographic recordings during mimicry of stimuli from the synthetic speech continuum /slt/-/split/. They found that "either there was a normal burst of muscle potential at the lip (indicating a /p/ gesture) or there was not [p. 842]." Since pairs of vocal responses that are topographically continuous, as well as pairs that are not, fail to show blending at stimulus-identification boundaries, it seems that the discreteness of an identification response is more a matter of response

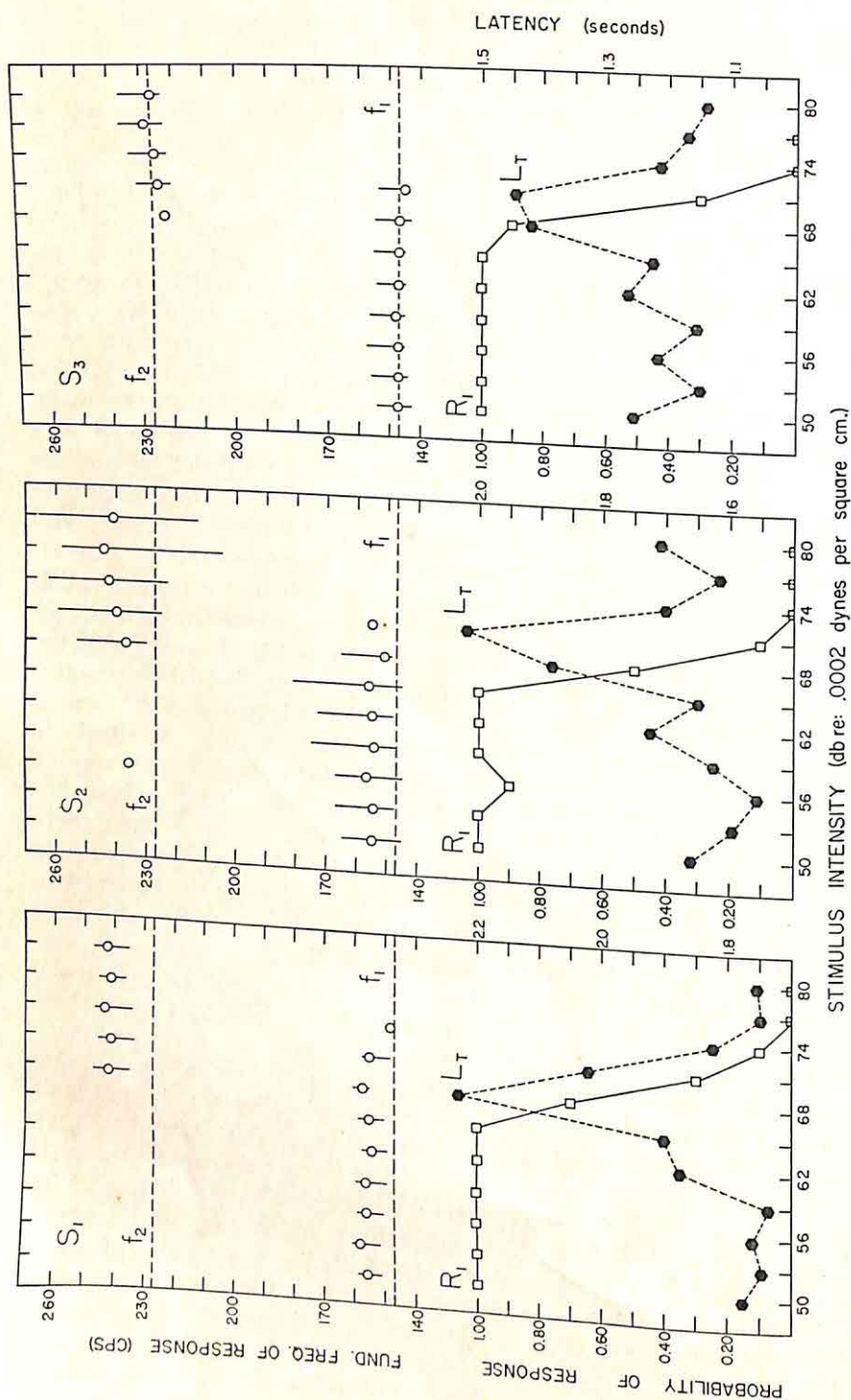


Fig. 13. Relations among response probability, latency, and topography in stimulus generalization. (Top: median frequency in cycles per second, circles, and the range, vertical lines, of high- and low-pitch responses as a function of stimulus intensity for each of three subjects. The dashed horizontal lines represent the vocal pitches previously differentiated. Bottom: response probabilities, squares, equal the ratio of the number of low-pitch responses emitted to the number of stimulus presentations, 10, at each intensity. The hexagons represent the average latency of high- and low-pitch responses to each stimulus intensity.)

stereotypy than it is of articulatory mechanics.

The general finding that the properties of identification, discrimination, latency, and response topography which are obtained with synthetic consonant stimuli are also replicated with synthetic vowels and with nonspeech stimuli is directly counter to the formulation of the motor theory of speech perception and the predictions which follow from it presented earlier in this article.

Psychophysical Functions for Speech Loudness and their Relation to Vocal Effort

Whereas the motor theory of speech perception propounded by Liberman and his co-workers would sort the perception of vowel quality and stress into one class (continuous articulation) and that of consonants into another (categorical articulation), Ladefoged (1959) excludes vowel quality from the domain of the motor theory and places stress in a common class with the consonants, despite the differences in the articulatory responses that produce them:

... vowel quality, nasality, and pitch are most simply related to an ordering of the acoustic properties of the stimulus, whereas most consonant qualities and features such as stress are ordered more simply in terms of articulatory activity [p. 406].

The evidence seen as favoring this hypothesis has been cited earlier, and some countervailing findings may now be described.

It was reported earlier that Lane et al. (1961) found that the *speaker's* estimates of his own vocal level, his autophonic output, grew as the 1.1 power of the sound pressure produced. Then Ladefoged and McKinney (1963) found that listeners' estimates of the loudness of speech sounds grew as the 1.2 power of the sound pressure

presented. These authors concluded that their listeners were, in fact, estimating the effort required to produce the sounds that were presented to them.¹² However, Lane et al. (1961) also obtained listener estimates of speech sounds, varying in autophonic level and in sound pressure level, and came up with a different finding: Listener estimates of the loudness of speech signals grew as the .7 power of the sound pressure—a function quite similar to the sone scale of pure-tone loudness.

By using magnetic recording techniques to decouple autophonic levels from their corresponding sound pressure levels, Lane (1962) was able to vary both parameters independently and to assess the contribution of each to estimates of speech loudness. When autophonic level was varied from whispering, at the one extreme, to shouting, at the other, the quality (spectrum) of the voice inevitably varied, but listener estimates of speech loudness remained invariant. These findings, replicated several times, show that estimates of the loudness of speech signals are not under the control of the effort involved in producing those signals but rather under the control of the sound pressure levels of the signals.

Further evidence that estimates of loudness are not estimates of effort is provided by experiments that require subjects to match their autophonic level to sounds of varying loudness and by complementary experiments that require subjects to decrease autophonic level, under increases in sidetone level, to maintain sidetone loudness constant.

¹² Whereas Ladefoged and McKinney (1963) contend that estimates of speech loudness are actually estimates of vocal effort, Warren (1962) has recently argued that estimates of vocal effort are actually estimates of loudness. For evidence that the motor theory works no better when stood on its head, see Lane (1963).

Both kinds of equal-sensation functions, which turn out to be reciprocal power functions, have an exponent (1.5) predicted by the ratio of the exponents of the autophonic scale and the loudness scale (Lane et al., 1961). In other words, not one but two disparate psychophysical functions are required to predict the behavior of the subject who listens to the loudnesses of speech signals and matches or regulates them by dint of his own vocal effort.¹³

The final experiment that was cited previously as providing support for the inference of mediating articulation in the perception of stress was conducted by Lehiste and Peterson (1959). These investigators found, it will be recalled, that when two vowels, generated with unequal effort, were presented to listeners at the same sound pressure level, the vowel generated with the greater effort was perceived as louder. The procedure employed in this experiment does not permit us to assess the relative contributions to the perception of loudness from sound pressure level and from effort cues when both are varying, as in natural language. Furthermore, any systematic difference whatsoever between the vowels in each pair may be misconstrued as determining perceived loudness if the listener is compelled to make loudness judgments in the absence of the acoustic cue normally associated with these judgments. For example, the vowel in each pair with the higher pitch may have been judged louder

simply because this was the only discriminable respect in which one vowel was in any way "greater" than the other.

In order to explore the relation between effort cues and sound pressure in determining judgments of vowel loudness, Lane (1962) recorded the phoneme /a/, produced at very low and very high levels, and played each of these at various intensities to listeners. The slope of the loudness functions, relating estimates to sound pressure levels, was invariant under wide differences in the levels at which the vowels were originally produced. In a subsequent experiment, Lane (1962) had speakers vocalize the phoneme /a/ at a wide variety of levels over a 40-decibel range and then played these vowels at a constant sound pressure level to listeners who were instructed to give magnitude estimates of vowel loudness. Under these conditions, with sound pressure level constant and effort or autophonic level varying, magnitude estimates did increase slightly as a function of the *original* sound pressure level; the "loudness" scale had a slope of less than 0.1. These findings indicate that the method of pair comparisons employed by Lehiste and Peterson (1959) gave a misleading image of the degree to which loudness judgments can be influenced by secondary properties of the vowels; ratio-scaling techniques reveal a small effect which may not be related to vocal effort.

SUMMARY AND CONCLUSIONS

Vocal responses may be differentiated in topography and then brought under discriminative control either by the language community or by an experimenter in the laboratory. If two discriminative stimuli, drawn from a physical continuum, control two differentiated vocal responses by an observer, and stimuli intermediate on the continuum are presented to him, certain characteristic relations obtain among the following de-

¹³ Contrary to a suggestion by Ladefoged and McKinney (1963, p. 460) this outcome cannot be attributed to a failure by the subjects to perceive the stimuli, some of them isolated vowels, as speech. Lane et al. (1961) cite five experiments by other investigators who used monosyllables or connected discourse—McMahon and Taylor (1963) have made it six—all of which confirm the basic disparity between the dynamics of speech and hearing.

pendent variables: stimulus identification, stimulus discrimination, identification latency, and response topography.

Identification of stimuli from a synthetic speech continuum, which ranges between two phonemes, constitutes a special case of this general paradigm in which (a) discrimination training and response differentiation have been conducted by the language community over a period of years, and (b) the identification response is formally similar to the acoustic stimulus. Under these conditions, it is tempting to speculate that the vocal responses, which in the language community generate the extreme stimuli on the continuum, are active in the perception of those stimuli when they are generated by a device or another person. This hypothesis has found some support in inferences drawn from the relations among the dependent variables enumerated above.

On the other hand, the inference that stimuli are perceived in terms of their articulatory correlates is unlikely to be formed in comparable experiments in which (a) discrimination training and response differentiation are accomplished quickly in the laboratory and (b) the identification response is not formally similar to the discriminative stimulus. Nevertheless, the same relations among identification, discrimination, latency, and topography are found in both kinds of experiments; those employing speech and those employing nonspeech continua.

Since basically the same conditioning and testing procedures are employed in both kinds of experiments—the speech perception studies are a special case of the general paradigm—and since the same relations among the four dependent variables are found in both, we conclude that these relations are attributable to the general paradigm for discriminative training and testing and that the postulation of a special perceptual mechanism for the discrimination of speech stimuli is unwarranted.

REFERENCES

- BASTIAN, J., & ABRAMSON, A. S. Identification and discrimination of phonemic vowel duration. *Journal of the Acoustical Society of America*, 1962, 34, 743-744. (Abstract)
- BASTIAN, J., EIMAS, P. D., & LIBERMAN, A. M. Identification and discrimination of a phonemic contrast induced by silent interval. *Journal of the Acoustical Society of America*, 1961, 33, 842. (Abstract)
- BEARE, A. C. Color name as a function of wavelength. *American Journal of Psychology*, 1963, 76, 248-256.
- CHISTOVICH, L. A., KLAAS, YU. A., & ALEKIN, R. O. The role of imitation in discrimination of sound series. *Voprosy Psikhologii*, 1961, 5, 173-182.
- COOPER, F. S. Spectrum analysis. *Journal of the Acoustical Society of America*, 1950, 22, 761-762.
- COOPER, F. S. Some instrumental aids to research on speech. In *Report of the Fourth Annual Round Table Meeting on Linguistics and Language Teaching*. Washington, D. C.: Institute of Language and Linguistics, Georgetown Univ., 1953.
- COOPER, F. S., LIBERMAN, A. M., & BORST, J. M. The interconversion of audible and visible patterns as a basis for research in the perception of speech. *Proceedings of the National Academy of Sciences*, 1951, 37, 318-325.
- COOPER, F. S., LIBERMAN, A. M., HARRIS, K. S., & GRUBB, P. M. Some input-output relations observed in experiments in the perception of speech. In *Proceedings of the Second International Congress on Cybernetics*. Namur, Belgium: International Association of Cybernetics, 1958.
- CROSS, D. V., & LANE, H. L. On the discriminative control of concurrent responses: The relations among response frequency, latency, and topography in auditory generalization. *Journal of the Experimental Analysis of Behavior*, 1962, 5, 487-496.
- CROSS, D. V., & LANE, H. L. An analysis of the relations between identification and discrimination functions for speech and nonspeech continua. Unpublished Report No. 05613-3-P, Behavior Analysis Laboratory, University of Michigan, 1964.
- CROSS, D. V., LANE, H. L., & SHEPPARD, W. C. Identification and discrimination functions for a visual continuum and their relation to the motor theory of speech perception. *Journal of Experimental Psychology*, 1965, 70, 63-74.
- DELATTRE, P. Les indices acoustique de la parole. Premier rapport. *Phonetica*, 1958, 2, 108-118, 210-251.
- DRAPER, M. H., LADEFOGED, P., & WHITERIDGE, D. Respiratory muscles in speech. *Journal of Speech and Hearing Research*, 1959, 2, 17-27.
- EIMAS, P. D. The relation between identification and discrimination along speech and nonspeech continua. *Language and Speech*, 1963, 6, 206-217.

- EKMAN, G. Contributions to the psychophysics of color vision. *Studium Generale*, 1963, 16, 54-64.
- FANT, C. G. M. Comments on 'A motor theory of speech perception' by A. M. Liberman, F. S. Cooper, K. S. Harris, & P. F. MacNeilage. In C. G. M. Fant (Ed.), *Proceedings of the Speech Communications Seminar*. Vol. 3. Unpublished report, Speech Transmission Laboratory, Royal Institute of Technology, Stockholm, 1963.
- FANT, G. *Acoustic theory of speech production*. The Hague: Mouton, 1960.
- FLANAGAN, J. L. Difference limen for the intensity of a vowel sound. *Journal of the Acoustical Society of America*, 1955, 27, 1223-1225.
- FONAGY, I. Elektrophysiologische Beiträge zur Akzent frage. *Phonetica*, 1958, 2, 12-58.
- FRY, D. B., ABRAMSON, A. S., EIMAS, P. D., & LIBERMAN, A. M. The identification and discrimination of synthetic vowels. *Language and Speech*, 1962, 5, 171-189.
- HARRIS, KATHERINE S., BASTIAN, J., & LIBERMAN, A. M. Mimicry and the perception of a phonemic contrast induced by silent interval: Electromyographic and acoustic measures. *Journal of the Acoustical Society of America*, 1961, 33, 842. (Abstract)
- HIRSCH, I. J. Auditory perception of temporal order. *Journal of the Acoustical Society of America*, 1959, 31, 759-767.
- HOCKETT, C. F. *A manual of phonology*. (Mem. 11) Indiana University Publications in Anthropology and Linguistics. Baltimore: Waverly Press, 1955. (*International Journal of American Linguistics*, 1955, 21(4), Part 1.)
- JAKOBSON, R., & HALLE, M. *Fundamentals of language*. Janua Linguarum, No. 1. The Hague: Mouton, 1956.
- JONES, L. A. The fundamental scale of pure hue and retinal sensibility to hue differences. *Journal of the Optical Society of America*, 1917, 1, 63-77.
- JONES, S. The accent in French—what is accent? *La Maitre Phonetique*, 1932, 40, 74-75.
- JOOS, M. Acoustic phonetics. *Language Monographs*, 1948, 24(23).
- KELLOGG, W. N. The time of judgment in psychometric measures. *American Journal of Psychology*, 1931, 43, 65-86.
- LADEFOGED, P. The perception of speech. In, *Mechanization of thought processes*. London: H. M. Stationery Office, 1959.
- LADEFOGED, P., DRAPER, M. H., & WHITTERIDGE, D. Syllables and stress. *Miscellanea Phonetica*, 1958, 3, 1-14.
- LADEFOGED, P., & MCKINNEY, N. P. Loudness, sound pressure, and subglottal pressure in speech. *Journal of the Acoustical Society of America*, 1963, 35, 454-460.
- LANE, H. L. Psychophysical parameters of vowel perception. *Psychological Monographs*, 1962, 76(44, Whole No. 563).
- LANE, H. L. The autophonic scale of voice level for congenitally deaf subjects. *Journal of Experimental Psychology*, 1963, 66, 328-331.
- LANE, H. L. Programmed learning of a second language. *International Review of Applied Linguistics*, 1965, 2, 249-301.
- LANE, H. L., CATANIA, A. C., & STEVENS, S. S. Voice level: Autophonic scale, perceived loudness, and effects of sidetone. *Journal of the Acoustical Society of America*, 1961, 33, 160-177.
- LANE, H. L., & KOPP, J. L. The effects of response dependent and independent reinforcement in extending stimulus control. *Psychological Record*, 1964, 14, 81-87.
- LANE, H. L., & MOORE, D. J. Reconditioning a consonant discrimination in an aphasic: An experimental case history. *Journal of Speech and Hearing Disorders*, 1962, 27, 232-243.
- LANE, H. L., & SCHNEIDER, B. A. Discriminative control of concurrent responses by the intensity, duration and relative onset time of auditory stimuli. Unpublished report, Behavior Analysis Laboratory, University of Michigan, 1963.
- LEHISTE, ILSE, & PETERSON, G. E. Vowel amplitude and phonemic stress in American English. *Journal of the Acoustical Society of America*, 1959, 31, 428-435.
- LENNEBERG, E. Understanding language without ability to speak: A case report. *Journal of Abnormal and Social Psychology*, 1962, 65, 419-425.
- LIBERMAN, A. M. Some results of research on speech perception. *Journal of the Acoustical Society of America*, 1957, 29, 117-123.
- LIBERMAN, A. M., COOPER, F. S., HARRIS, KATHERINE S., & MACNEILAGE, P. F. A motor theory of speech perception. In C. G. M. Fant (Ed.), *Proceedings of the Speech Communications Seminar*. Unpublished report, Speech Transmission Laboratories, Royal Institute of Technology, Stockholm, 1963.
- LIBERMAN, A. M., DELATTRE, P., & COOPER, F. S. The role of selected stimulus vari-

- ables in the perception of the unvoiced stop consonants. *American Journal of Psychology*, 1952, 65, 497-516.
- LIBERMAN, A. M., DELATTRE, P. C., & COOPER, F. S. Some cues for the distinction between voiced and voiceless stops in initial position. *Language and Speech*, 1958, 1, 153-157.
- LIBERMAN, A. M., DELATTRE, P. C., COOPER, F. S., & GERSTMAN, L. J. The role of consonant-vowel transitions in the perception of stop and nasal consonants. *Psychological Monographs*, 1954, 68(8, Whole No. 379).
- LIBERMAN, A. M., HARRIS, KATHERINE S., EIMAS, P., LISKE, L., & BASTIAN, J. An effect of learning on speech perception: The discrimination of durations of silence with and without phonemic significance. *Language and Speech*, 1961, 4, 175-195.
- LIBERMAN, A. M., HARRIS, KATHERINE S., HOFFMAN, H. S., & GRIFFITH, B. C. The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*, 1957, 54, 358-368.
- LIBERMAN, A. M., HARRIS, KATHERINE S., KINNEY, JOANN, & LANE, H. L. The discrimination of relative onset time of the components of certain speech and non-speech patterns. *Journal of Experimental Psychology*, 1961, 61, 379-388.
- LIBERMAN, A. M., STEVENS, K. N., & OHMAN, S. Identification and discrimination of synthetic speech sounds. Unpublished Report No. STL-QPSR-2, Royal Institute of Technology, Stockholm, 1963.
- LISKE, L., COOPER, F. S., & LIBERMAN, A. M. The uses of experiment in language description. *Word*, 1962, 18, 82-106.
- McMAHON, L., & TAYLOR, J. Direct comparisons of the autophonic scale and the loudness scale. Paper read at Acoustical Society of America, New York, 1963.
- MILLER, G. A. The magical number seven, plus or minus two. *Psychological Review*, 1956, 63, 81-97.
- PETERSON, G. E., & BARNEY, H. L. Control methods used in a study of the vowels. *Journal of the Acoustical Society of America*, 1952, 24, 175-184.
- PRINS, D. Relations among specific articulatory deviations and responses to a clinical measure of sound discrimination ability. *Journal of Speech and Hearing Disorders*, 1963, 28, 382-388.
- STEVENS, K. N. Toward a model for speech recognition. *Journal of the Acoustical Society of America*, 1960, 32, 47-55.
- STEVENS, K. N., OHMAN, S. E. G., & LIBERMAN, A. M. Identification and discrimination of rounded and unrounded vowels. Paper read at Acoustical Society of America, Ann Arbor, 1963.
- STUDDERT-KENNEDY, M., LIBERMAN, A. M., & STEVENS, K. N. Reaction time to synthetic stop consonants and vowels at phoneme centers and at phoneme boundaries. Paper read at Acoustical Society of America, Ann Arbor, 1963.
- THURLOW, W. R. Perception of low auditory pitch. *Journal of Experimental Psychology*, 1963, 70, 461-470.
- WADSWORTH, G. P., & BRYAN, J. G. *Introduction to probability and random variables*. New York: McGraw-Hill, 1960.
- WARREN, R. M. Are "autophonic" judgments based on loudness? *American Journal of Psychology*, 1962, 75, 452-456.

(Received April 3, 1964)

ON THE COMBINATION OF DRIVE AND INCENTIVE MOTIVATION

ROGER W. BLACK¹

State University of Iowa

Hull assumed that the theoretical variables D and K combine in a multiplicative manner while Spence has suggested that the relation is additive. The great majority of studies dealing with this issue have supported Spence's view in that they have obtained no $D \times K$ interaction. Experiments which have involved "0" values of deprivation and reward, however, have reported such an interaction. The present paper is concerned with describing the implications for both types of experiments of Spence's interpretation of K when that interpretation is coupled with the view that K is a function of level of deprivation. This analysis appears to successfully integrate the results from studies concerned with the combination of D and K . In addition, it derives the findings of recent experiments concerned with the possibility that acquisition-drive level may have an historical or perseverating effect on performance under subsequent conditions of drive.

The manner in which the theoretical variables incentive motivation (K) and drive (D) combine to produce reaction potential (E) has been the subject of considerable experimental interest in recent years. One obvious source of this interest lies in the fact that both Hull (1952) and Spence (1956) have advanced explicit, but different, formulations of the relationship. Thus, according to Hull, K and D combined multiplicatively, while Spence has suggested that the relationship is additive. A second factor that has facilitated such research is the apparent ease with which the experimental problem can be defined. Thus, when suitable operations for manipulating D and K are selected, the factorial design appears to provide an ideal experimental test of these alternative assumptions. For such investigations, typically involving two or more levels of both D and K , Hull's multiplicative formulation leads to the expectation of a significant D

$\times K$ interaction while Spence's additive assumption predicts no interaction. Stated graphically, if performance is plotted as a function of D with K as the parameter, Spence's additive formulation predicts parallel curves while Hull's multiplicative assumption requires that the functions diverge.

Theoretical Relationship between D and K

As noted later, the bulk of evidence from "typical" factorial studies has supported Spence's position. Nevertheless, several experiments do report an interaction between time of deprivation (T_d) and magnitude of reward and have been interpreted as demonstrating a nonadditive relationship between D and K . It must be noted, however, that although Spence has assumed that the combination of the theoretical variables D and K is additive, his treatment of K is not such as to preclude the possibility of an interaction between the experimental variables employed to manipulate these theoretical variables. The present

¹ The writer is indebted to Kenneth W. Spence who read the present manuscript and provided helpful comments and suggestions.

paper is concerned with describing the conditions under which an interaction of these empirical variables is not only compatible with Spence's interpretation of these variables but is, indeed, implied by it.

According to Spence (1956), K is the molar theoretical variable that summarizes the motivational consequences of the $r_g - s_g$ mechanism which is assumed to consist of implicit components of the overt consummatory response (R_G) that become conditioned to the stimulus situation in which the instrumental response is appetitively reinforced.² The strength of the $r_g - s_g$ mechanism depends, in turn, upon at least two variables. First, since it results from classical conditioning to goal-box cues, the strength of $r_g - s_g$ will be an increasing function of the frequency and duration with which R_G has been elicited in the presence of those cues. Thus, K is assumed to be an increasing function of the number of rewarded trials and approaches an asymptote determined, in part, by the magnitude of reward. Second, since the strength of $r_g - s_g$ is assumed to

reflect the vigor of the overt consummatory response (R_G) of which it is a component, any variable which contributes to the vigor of R_G will also affect the theoretical value of K .

A possibility implicit in this formulation which has not been previously elaborated is that the procedures employed to produce D (e.g., Td) are also determinants of K in instrumental reward conditioning. There are several bases on which such a relationship might be expected. Thus, since the consummatory response (R_G) is an overt, skeletal response, its strength or vigor should vary directly with D in the manner typical of other such responses. It is well known, for example, that the amount of food or water which a subject will consume varies directly with length of deprivation, and Stellar and Hill (1952) have concluded that this type of consummatory behavior—at least in the case of thirst—provides the most accurate, available index of drive level. Similarly, Snyder (1962) has reported that the rate at which rats licked for a saccharine reward is an increasing function of D , although Stellar and Hill (1952) and Collier (1962) have found licking rate to be relatively independent of D . Even in cases in which the vigor of R_G is not strongly affected by Td, however, Td may still act as a determinant of K , both through its facilitation of r_g directly and through its effect on the height of the gradient of generalization of r_g from the goal box to the runway, etc.

Level of deprivation may also contribute to K by modifying the "palatability" of incentives or the threshold for the occurrence of consummatory behavior. Thus, there may be no food incentive sufficiently "palatable" to evoke a consummatory response in a "completely" satiated subject while as satiation decreases, increasingly less

² The $r_g - s_g$ mechanism is an abstract or theoretical concept which, as such, does not require explicit speculation or specification as to its physiological nature or locus of occurrence. It is true, however, that r_g has conventionally been considered to consist of peripheral responses such as salivation, stomach contractions, etc. Miller (1963), for example, has recently presented an interpretation of reinforcement which resembles, in part, Spence's treatment of the $r_g - s_g$ mechanism. Miller notes, however, that his analysis specifically "does not limit itself to incentives based on the conditioning of peripheral consummatory responses . . . [p. 97]." Actually, the nature of Spence's treatment of $r_g - s_g$ does not compel him to limit that mechanism to peripheral responses nor has he so limited himself by explicit assumption. The present paper also assumes that r_g 's may be both peripheral or central in nature and that the argument developed here is equally applicable to either type of implicit response.

palatable incentives may exceed the threshold for R_G . Evidence of such an influence of Td on the palatability or incentive value of rewards is reported by Stabler (1962) who found that runway performance was a decreasing function of sucrose concentration under low Td, but an increasing function of concentration under high Td. One possible interpretation of this finding is, of course, that the incentive value (or, K) associated with different sweetnesses depended on Td. Further corroboration of this finding that the palatability or incentive value of a reward varies with Td was reported by Snyder (1962) who found that the reinforcement value of a presumably non-nutritive reward (saccharine) was, like that of sucrose, a function of Td. Unlike the result with sucrose, however, performance under high Td was a decreasing function of saccharine concentration, while with low Td performance increased with concentration.

Empirical Relationship between Deprivation and Reward Magnitude

The possibility that K is partially determined by the experimental procedures producing D is, therefore, not without empirical support. In addition, this hypothesis, together with the assumption that K and D combine in an additive fashion, currently appears to provide the most adequate approximation to the relevant empirical findings. For example, a recent series of studies by Seward and his associates have been interpreted as supporting the view that D and K interact—at least when extreme values of these variables are employed. Thus, Seward, Shea, and Elkind (1958) factorially combined two levels of reward with two levels of deprivation in a runway-conditioning situation. Magnitude of reward was either 0 or 1.00 gram(s) of wet mash while D was varied by testing the sub-

jects either immediately before or after administration of their total daily ration, thus providing one group which was 23-hours hungry and one which was 0-hours hungry and, presumably, satiated. Letting lowercase letters indicate low, and capital letters, high values of the theoretical variables D and K , the groups were: DK, Dk, dK, and dk. For the latter three groups, little evidence of the acquisition of the instrumental running response was found and only slight differences between the groups occurred, although performance was somewhat better for Dk. Group DK, on the other hand, produced a definite acquisition curve and ran considerably faster than the other groups. The fact that Groups dk, dK, and Dk were about equal and inferior to DK gave rise to a significant $D \times K$ interaction, since the difference between, for example, DK and Dk was thus necessarily much larger than that between dK and dk.

While Seward et al. (1958) interpreted these results as suggesting that Spence's additive assumption applies only for intermediate levels of deprivation and reward, their data are actually quite compatible with the assumptions that D and K combine additively while K is a function of both reward magnitude and deprivation level. Thus, if the 0-hours deprivation groups were largely or completely satiated, it is unlikely that the subjects in those groups consumed much or any food during the 15–30 seconds they remained in the baited goal box. In this event, since R_G was not evoked, K would develop no more than for the subjects in the "zero reward" groups and hence, Group dK should show performance essentially identical to that of Group dk while neither should improve with training, no "reinforcement" being administered. Group Dk should similarly fail to improve with training, but

would be expected to exhibit a higher level of performance than dK or dk because of its higher level of D . Finally, Group DK should improve with training as a result of its acquisition of K and should be significantly superior to any of the other groups, since these groups acquire no K . These expectations were clearly confirmed by the Seward et al. study, and a similar interpretation may be placed on related studies by Seward and Proctor (1960) and Seward, Shea, and Davenport (1960) mentioned below.

While the assumption that K is partially determined by D can apparently account for the interaction found in those studies which involve zero values of deprivation and reward, in the more numerous experiments which employ no zero values of these variables, no $D \times K$ interaction is typically reported. Thus, with D defined by degree of food deprivation and K by magnitude of food reward, no interaction was found by Logan (1960), Reynolds and Pavlik (1960), or Weiss (1960). When D was manipulated by differential reduction in the subject's body weight, no significant interaction was found with sucrose concentration (Brush, Goodrich, Teghtsoonian, & Eisman, 1961) nor with magnitude of food reward (Yarczower, Freygold, & Blum, 1962). Snyder (1962) varied deprivation time for food and water and concentration of saccharine and also failed to obtain the interaction predicted by Hull. On the other hand, Ehrenfreud and Badia (1962) did obtain an interaction between D , as defined by reduction in body weight and amount of food reward, but the interaction was opposite to that predicted by Hull—i.e., curves for groups trained under small versus large rewards actually converged with increasing D —a result

these experimenters interpret as a case of "exponential addition" (Spence, 1953). Finally, while the above experiments all involved runway conditioning, similar results were obtained by Hulicka (1960a, 1960b) for a bar-pressing response.

It is apparent, then, that in the typical experiment employing no zero levels of reward or deprivation, a Deprivation \times Reward interaction is not obtained. The present hypothesis, however, assumes that K varies with deprivation level as well as reward magnitude—an assumption which may appear to imply that a Deprivation \times Reward interaction should be expected generally. In the type of experimental situation in which no Deprivation \times Reward interaction has been found, however, none is, in fact, required by the present analysis. For expository purposes, for example, consider a 2×2 factorial design involving "low" and "high" values of both reward magnitude and deprivation level. Assuming that K is a function of both deprivation and reward, that K and D combine additively, and that a numerical weight of 1, 2, or 3 may be assigned to the low, medium, and high values of each of these variables, the theoretical expectations for such an experiment are represented in Table 1. Note that in each cell D has a value of either 1 or 3, since this variable is independent of reward magnitude, while K assumes values of 1-3, reflecting both magnitude of reward and deprivation level. The total in each cell represents the additive combination of D and K , and the marginal figures indicate differences between cells adjacent in rows or columns. Inspection of these marginal differences indicates clearly that no interaction of level of deprivation and reward magnitude is predicted. Moreover, this analysis can readily be generalized to more

TABLE 1

IMPLICATIONS OF THE PRESENT ANALYSIS
REGARDING THE THEORETICAL COM-
BINATION OF DEPRIVATION
AND REWARD

Level of deprivation	Magnitude of reward		
	Low	High	Diff.
Low	$K=1$ +	$K=2$ +	1
	$D=\frac{1}{2}$	$D=\frac{1}{3}$	
High	$K=2$ +	$K=3$ +	1
	$D=\frac{3}{5}$	$D=\frac{3}{6}$	
Diff.	3	3	

complex factorial designs. Thus, the assumptions that D and K combine additively, and that K is a function of T_d can account for the empirical evidence from those studies involving zero levels of the experimental variables as well as from those not employing such values.

Hull's view that K and D combine in a multiplicative fashion, on the other hand, does not fare well in light of the evidence reviewed above. Thus, since Hull did not specifically identify the molar variable K with the $r_g - s_g$ mechanism, the theoretical considerations outlined in connection with the studies of Seward and his associates are not directly applicable to the interpretation Hull would presumably place upon their data. Nevertheless, Hull's formulation does appear capable of deriving the results of the Seward et al. (1958) study described earlier. According to Hull, the associative (S^E-R), motivational (D), and incentive (K) factors all combine in a multiplicative fashion to determine reaction potential (S^E-R). Thus a zero value of any one

of these variables would reduce S^E-R to zero and result in performance at an "operant" level. Since Seward and his associates' experiment involved three groups (dk , dK , and Dk) which had at least one of these variables approaching zero value, Hull's theory would predict that the performance of these groups would be about equal, would not improve with training, and would be inferior to Group DK for which neither D nor K was zero. These predictions are, of course, similar to those deriving from the present assumptions and were confirmed by the results of Seward et al.

The experimental result which is of primary embarrassment to Hull's formulation is the fact that an interaction between reward magnitude and T_d is typically observed *only* when a zero value of one of the variables is employed. For example, Seward, Shea, and Davenport (1960) employed three levels of both drive and reward in a factorially designed, runway-conditioning experiment. Finding a Deprivation \times Reward interaction of minimal statistical reliability for the entire 3×3 table, these experimenters examined the nine simple 2×2 interactions which their design provided. Of the three simple interactions that proved significant, two involved both zero reward and the prefeeding of the subject its entire daily diet (1 hour access to food) just prior to training (i.e., 0 T_d), while the third significant interaction also involved the 0-hours deprivation group. Although the interaction observed when T_d or reward was zero is compatible with both Hull's and the present analysis, the failure to obtain an interaction in the six comparisons not involving zero values conforms only to the implications of the latter view and not to Hull's multiplicative assumption. Furthermore, the implications of Hull's formulation also

receive no support from the more conventional factorial studies since these experiments—with considerable consistency—have not observed the interaction of T_d and reward magnitude required by his position. Thus, Table 2 represents the predictions deriving from Hull's assumptions that D and K combine multiplicatively, that K is independent of T_d , and, for present purposes, that a value of 1 or 3 can be assigned to the high and low values of each of these variables. Examination of the marginal figures indicating differences in adjacent cells clearly demonstrates that a Deprivation \times Reward interaction is expected. Moreover, even when Hull's multiplicative assumption is coupled with the view that K is a function of T_d , a Deprivation \times Reward interaction must still be predicted—a prediction which, as noted earlier, has repeatedly failed confirmation.

The Effect of Acquisition Drive on Subsequent Performance

A further problem of considerable theoretical interest and to which the present hypothesis readily extends is that of whether the level of D present during an initial training session in instrumental reward conditioning affects performance in a subsequent testing period under altered, but equated, D conditions—i.e., whether D has an "historical effect" on performance. Early Skinner-box studies concerned with this question tended to find no such perseverating or historical effect of D (e.g., Finan, 1940; Kendler, 1945), but more recent runway studies employing careful procedures to insure the adequate equating of D prior to testing have rather consistently found an effect (e.g., Barry, 1958; Brush, Goodrich, Teghtsoonian, & Eisman, 1963; Campbell & Kraeling, 1954; Lewis & Cotton, 1957). If D present

TABLE 2
IMPLICATIONS OF HULL'S ASSUMPTIONS
REGARDING THE THEORETICAL COM-
BINATION OF DEPRIVATION
AND REWARD

Level of deprivation	Magnitude of reward		
	Low	High	Diff.
Low	$K=1$ \times $D=\frac{1}{1}$	$K=3$ \times $D=\frac{1}{3}$	2
High	$K=1$ \times $D=\frac{3}{3}$	$K=3$ \times $D=\frac{3}{9}$	6
Diff.	2	6	

during the acquisition of an instrumental response does, in fact, affect the performance of that response under subsequent, equated conditions of D , one interpretation might be that the associative factor (H) is a function of D . Nevertheless, since H , once acquired, is usually considered to be relatively stable or permanent, this assumption leads to the expectation of a persistent influence of acquisition D on subsequent performance. The usual finding, however, has been that the perseverating effect of acquisition D , when obtained, is quite transitory. Thus, if a group trained under high D and tested under low D is compared with a group trained and tested under low D , the superiority in performance of the former group typically disappears after the first few trials, thus suggesting that some historical factor other than H has been affected by acquisition D .

The hypothesis proposed in the present paper—that deprivation level is a determinant of K —also leads to the expectation of an historical effect of D

in instrumental reward conditioning. Thus, since it is presently assumed that K is an historical variable which is a function of the level of T_d under which it is acquired, those subjects trained under high D will enter the testing phase of such an experiment with a higher level of K than the subjects trained under low D . Due to the reduction in D , however, the strength of R_G for the initially high D group will be lessened at the beginning of the test phase. Further, since the strength or vigor of r_g is assumed to reflect that of R_G , $r_g - s_g$ and, hence, K will also undergo rapid reduction following the shift to lowered D . The historical effect of acquisition D , therefore, while required by the present hypothesis, would also be expected to be quite transitory. Thus, the hypothesis that K is a function of T_d not only constitutes an alternative to the view that H is a function of D in instrumental reward conditioning, but appears, in addition, to provide a more adequate description of the relevant empirical evidence.

REFERENCES

- BARRY, H., III. Effects of strength of drive on learning and extinction. *Journal of Experimental Psychology*, 1958, 55, 473-481.
- BRUSH, F. R., GOODRICH, K. P., TEGHTSOONIAN, R., & EISMAN, E. H. Running speed as a function of deprivation condition and concentration of sucrose incentive. *Psychological Reports*, 1961, 9, 627-634.
- BRUSH, F. R., GOODRICH, K. P., TEGHTSOONIAN, R., & EISMAN, E. H. Dependence of learning (habit) in the runway on deprivation under three levels of sucrose incentive. *Psychological Reports*, 1963, 12, 375-384.
- CAMPBELL, B. A., & KRAELING, D. Response strength as a function of drive level during training and extinction. *Journal of Comparative and Physiological Psychology*, 1954, 47, 101-103.
- COLLIER, G. Consummatory and instrumental responding as functions of deprivation. *Journal of Experimental Psychology*, 1962, 64, 410-414.
- EHRENFREUND, D., & BADIA, P. Response strength as a function of drive level and pre- and postshift incentive magnitude. *Journal of Experimental Psychology*, 1962, 63, 468-471.
- FINAN, J. L. Quantitative studies in motivation: I. Strength of conditioning in rats under varying degrees of hunger. *Journal of Comparative Psychology*, 1940, 29, 119-134.
- HULICKA, I. M. Additive versus multiplicative combination of drive and incentive. *Psychological Reports*, 1960, 6, 403-409. (a)
- HULICKA, I. M. Combination of drive and incentive. *Quarterly Journal of Experimental Psychology*, 1960, 12, 185-189. (b)
- HULL, C. L. *A behavior system*. New Haven: Yale Univ. Press, 1952.
- KENDLER, H. H. Drive interaction: II. Experimental analysis of the role of drive in learning theory. *Journal of Experimental Psychology*, 1945, 35, 188-198.
- LEWIS, D. J., & COTTON, J. W. Learning and performance as a function of drive strength during acquisition and extinction. *Journal of Comparative and Physiological Psychology*, 1957, 50, 189-194.
- LOGAN, F. A. *Incentive*. New Haven: Yale Univ. Press, 1960.
- MILLER, N. E. Some reflections on the law of effect produce a new alternative to drive reduction. In M. R. Jones (Ed.), *Nebraska symposium on motivation: 1963*. Lincoln: Univ. Nebraska Press, 1963. Pp. 94-98.
- REYNOLDS, W. F., & PAVLIK, W. B. Running speed as a function of deprivation period and reward magnitude. *Journal of Comparative and Physiological Psychology*, 1960, 53, 615-618.
- SEWARD, J. P., & PROCTOR, D. M. Performance as a function of drive, reward, and habit strength. *American Journal of Psychology*, 1960, 73, 448-453.
- SEWARD, J. P., SHEA, R. A., & DAVENPORT, R. H. Further evidence for the interaction of drive and reward. *American Journal of Psychology*, 1960, 73, 370-379.
- SEWARD, J. P., SHEA, R. A., & ELKIND, D. Evidence for the interaction of drive and reward. *American Journal of Psychology*, 1958, 71, 404-407.
- SNYDER, H. L. Saccharine concentration and deprivation as determinants of instrumental and consummatory response strengths. *Journal of Experimental Psychology*, 1962, 63, 610-615.

- SPENCE, K. W. Current interpretations of learning data and some recent developments in stimulus-response theory. In, *Learning theory, personality theory, and clinical research: The Kentucky symposium*. New York: Wiley, 1953. P. 20.
- SPENCE, K. W. *Behavior theory and conditioning*. New Haven: Yale Univer. Press, 1956.
- STABLER, J. R. Performance in instrumental conditioning as a joint function of time of deprivation and sucrose concentration. *Journal of Experimental Psychology*, 1962, 63, 248-253.
- STELLAR, E., & HILL, J. H. The rat's rate of drinking as a function of water deprivation. *Journal of Comparative and Physiological Psychology*, 1952, 45, 96-102.
- WEISS, R. F. Deprivation and reward magnitude effects on speed throughout the goal gradient. *Journal of Experimental Psychology*, 1960, 60, 384-390.
- YARCZOWER, M., FREYGOLD, K., & BLUM, N. Effect of amount, percentage of reinforcement and deprivation condition on runway time. *Psychological Reports*, 1962, 11, 406.

(Received July 24, 1964)

THEORETICAL NOTES

COMMENTS ON MANDLER'S "FROM ASSOCIATION TO STRUCTURE"

JOHN JUNG¹

California State College, Long Beach

A critical analysis of some of the evidence from transfer studies cited by Mandler (1962) as supportive of a concept of cognitive structures which develop from gradual associationistic processes revealed inadequacies in his formulation. Furthermore, recent studies on the effects of degree of 1st-list learning (with controls for nonspecific transfer) and response M on transfer in the A-B, A-C paradigm also are difficult to handle by the concept of structure. Nor is it clear how older findings of the effects of overlearning of interfering lists in retroactive and proactive inhibition paradigms can be explained by Mandler. A return to associative concepts such as response competition, availability, and differentiation was suggested, and it was shown how such concepts can account for the above findings which are incompatible with the concept of structure. It was suggested that structure may be useful in other areas but apparently not in transfer of rote learning.

Mandler (1962) has attempted to account for the phenomena of learning sets and the facts of overlearning in transfer of training in terms of cognitive *structures* which develop from associationistic processes. Such structures refer to central analogic representations of highly integrated response sequences which can function independently of the overt responses. According to Mandler, it is by means of such structures that animals can form learning sets, develop observing responses, and solve successive discrimination-learning problems more rapidly. For the same reason, human subjects learning verbal materials in the A-B, A-C (hereafter A-C) paradigm which typically leads to negative transfer can eliminate competing B responses from the first list during the learning of the second or transfer list if the first list was overlearned.

The present comments are directed at Mandler's interpretation of the facts of overlearning or degree of first-list learning in transfer tasks. Mandler's survey of relevant animal and human studies indicated that negative transfer first increases, then decreases, and in some cases

becomes positive transfer as a function of overlearning of the first list. It was concluded on this evidence of a U function between overlearning and negative transfer that structures are formed gradually out of associations.

Several objections to the necessity of invoking the concept of structures to account for such evidence will be made. In the first place, the soundness and, in some cases, the appropriateness of the findings on which Mandler based his formulation can be questioned. In all of the animal studies cited, as Mandler himself noted (pp. 422, 425, 426), no control group for assessing warm-up and learning-to-learn was included. How much of the subsequent decrease in negative transfer as overlearning increased in those studies is attributable to specific and how much to nonspecific sources of transfer is hard to say.

Furthermore, discrimination-learning situations such as those used in the animal studies cited are quite dissimilar from the A-C transfer task, but Mandler lumps them together since they both apparently involve the learning of two distinct responses to the same stimulus. Closer examination reveals that, whereas in a dis-

¹ Now at York University, Toronto.

crimination-reversal task two stimulus cues (e.g., black and white) and two responses (e.g., right and left turn) are available on *every* trial, in the A-C paradigm the subject receives only one response paired with each stimulus during the first list, and on the second list he receives a new second set of responses, one paired with each of the same stimuli. The animal-learning situations cited by Mandler are more closely related to the A-B, A-Br (hereafter A-Br) paradigm in which the same stimuli and responses appear on both tasks but in different pairings.

With respect to the three human studies using the A-C paradigm cited by Mandler, only one (Mandler & Heinemann, 1956) contained an A-B, C-D control condition to assess warm-up and learning-to-learn. In addition to this absence, the Siipola and Israel (1933) study, although showing positive transfer with overlearning, involved an A-Br rather than an A-C paradigm. Positive transfer has been found when low meaningfulness (M) or difficult responses such as those used by Siipola and Israel are employed in the A-Br paradigm (Mandler & Heinemann, 1956; Merikle & Battig, 1963). However, substantial negative transfer can be obtained with this paradigm when high-M responses are used (Besch & Reynolds, 1958; Jung, 1962; Merikle & Battig, 1963; Porter & Duncan, 1953; Postman, 1962b; Twedt & Underwood, 1959). Transfer of response learning rather than the formation of structures can account for the positive transfer obtained with this paradigm with low-M responses.

Although evidence from transfer studies cited by Mandler is weak support for his concept, recent findings which were not available at the time of Mandler's paper but which include the necessary control for nonspecific transfer do suggest that transfer under the A-C paradigm may be a U function of degree of first-list learning to a small extent (Jung, 1962; Postman, 1962b). Although neither study reported significant differences in negative transfer as a function of first-list overlearning, Postman found a U-shaped trend, and Jung, using only two

degrees of first-list overlearning, found trends toward less negative transfer with overlearning during the last half of the transfer task. No positive transfer was obtained as in studies lacking the proper control condition; this suggests that competition between the associations of the two lists is still operative with overlearning.

Secondly, there are certain established findings regarding the effects of overlearning of the interfering task in retroactive and proactive inhibition (RI and PI, respectively) studies which appear to be contrary to a notion of cognitive structure. With PI, a number of studies have found that increased first-list learning leads to enhanced PI (e.g., Atwood, 1953; Postman & Riley, 1959). The concept of structure should predict *less* PI, as the first list is overlearned. In the case of RI, findings (e.g., Briggs, 1957; Thune & Underwood, 1943; Underwood, 1945) indicate that the greater the degree of learning on the interpolated list, the more RI, although in some cases it decreases with high overlearning but not to the extent of zero RI which a concept of structure would predict ultimately. Thus, in both PI and RI paradigms, overlearning of the interfering list relative to the degree of learning of the list-to-be-recalled results in greater interference rather than a reduction of negative effects as implied in the concept of structure.

Third, recent experiments (Jung, 1963; Merikle & Battig, 1963; Spence, 1963) demonstrating the effects of response M on transfer in the A-C paradigm also provide difficulty for Mandler. If the writer understands the concept of structure correctly, structure is a direct function of M. In the A-C paradigm, therefore, less negative transfer should result with responses of high M (or structure). Small but consistently greater negative transfer has been found with high than with low M when trigrams were used (Jung, 1963) or when words and trigrams were used (Merikle & Battig, 1963). A different method of assessing the effects of first-list structure on A-C transfer was employed in one of Spence's (1963) experiments. She employed a first list which

contained pairs of words which were strong associates of each other; on the transfer list she used the same stimuli from the first list but paired them with a different set of words which were weak associates of the stimuli. Spence maintained that the first-list associates which she used undoubtedly had more structure than any practical amount of laboratory overlearning could provide since the strong preexisting associations used in that list had been formed over a subject's lifetime.

According to the concept of structure, the first list of strong associates should not interfere with the learning of the transfer list containing weak associates. However, Spence reported that negative transfer was obtained even under these conditions.

Thus several recent studies of the effects of response M and high preexisting associative first-list pairs on transfer in the A-C paradigm provide evidence clearly opposed to the concept of structure.

From Structure Back to Association

As Mandler noted, one should "... then outline the major variables apparently involved in the phenomena, and finally *examine the extent to which associative factors may account for them, before invoking analogic structures as operative in these situations* [p. 419; *italics mine*]." An attempt will now be made to account for the above evidence which appears to be incompatible with a concept of structure by means of associative concepts.

The general findings of increases, then decreases, in negative transfer in the A-C paradigm can be accounted for by the associative concepts of *response competition*, *response availability*, and *response differentiation*. In this paradigm where dissimilar responses are paired with the same stimuli on the two lists, the only source of specific transfer is that of associative learning (Jung, 1963, p. 378), i.e., associations formed on the first list interfere with the formation of new associations to the same stimuli on the second list, hence, negative transfer.

Barnes and Underwood (1959) and

Postman (1962a, 1962b) demonstrated that after the learning of the second list in the A-C paradigm, the first-list responses were less available than the second-list responses as measured by a modified free-recall task. Furthermore, Barnes and Underwood found that this difference in availability increased with higher degrees of second-list learning. This diminished response availability suggested to Barnes and Underwood that the competing first-list associations underwent extinction during second-list learning.

By analogy, increasing the degree of *first-list* learning for a constant degree of second-list learning should *increase* availability of competing first-list associations during second-list learning, offsetting to some extent the extinction which such associations also undergo. Indirect support for this hypothesis can be found in the fact mentioned earlier that increased PI (poorer retention of the second list) results as a function of overlearning of the first list.

The increase in negative transfer observed with moderate overlearning of the first list (Jung, 1962; Postman, 1962b) may be due to the fact that the interfering first-list associations are simply more available during second-list learnings. But, as has already been pointed out, high overlearning of the first list leads to a decrease in A-C negative transfer. Postman (1962b) accounted for this failure of high overlearning to lead to even greater negative transfer by means of the differentiation which had developed. When two lists are practiced to an equal degree, the differentiation between the lists is minimal; however, as the amount of practice on one list relative to that on the other increases, differentiation may be said to occur. Such is the case with high overlearning of the first list in the A-C paradigm; such differentiation reduces the competition between responses from the two lists, thus producing less negative transfer.

One difficulty with this explanation is that while differentiation may lead to decreases in overt intrusions (Thune & Underwood, 1943; Underwood, 1945), it does not necessarily lead to reduced in-

interference in the RI task. While Underwood may be correct in maintaining that differentiation does not reduce RI (or PI), this is not to rule out the possibility that differentiation may be a potent aid in reducing response competition and negative transfer during the *original learning of the second list*. According to Underwood, differentiation dissipates over time; since the retention of the first list is measured a long time after the learning of the second list in the RI paradigm, the ameliorative effects of differentiation on retention may be small or absent due to its dissipation during the long interval between interpolated learning and a retention test of the first list.

Turning to the effects of overlearning in RI and PI studies which appeared incompatible with the concept of structure, response availability seems to be a useful concept. It will be recalled that more PI is generally found, the higher the first-list learning, and that more RI is generally found, the higher the second-list learning. Both cases could be due to the increased availability of associations which are incompatible with those required in the less well-learned list-to-be-recalled. As pointed out earlier, due to its transitory nature, any differentiation which also develops with overlearning may not reduce negative effects in retention (RI and PI) although it may assist original learning of the second list.

Finally, how can the findings of greater negative transfer with high-M responses be accounted for by associative concepts? If it can be assumed that response availability is a direct function of response M, then the finding of higher negative transfer with high M is not surprising. That is, associations formed on the first list may either be forgotten or extinguished faster if they are of low M; thus they would be less available to compete with the new associations required on the second list. No direct evidence on how response M affects response availability is available to the writer's knowledge. However, for a wide range of M from nonsense syllables to words, the assumption that availability is a direct function of M might be based on the letter-se-

quence gradient of extraexperimental interference (Underwood & Postman, 1960). According to this notion, letter sequences contained in experimental materials are subject to such extraexperimental interference (and forgetting) to the extent that they differ from those sequences occurring in the natural language. Thus, low-M materials would suffer the most interference from this extraexperimental source, and their availability would be diminished accordingly.

Consequently, negative transfer may be less with low-M responses simply because low-M associations are less available during second-list learning. On the other hand, a concept such as Mandler's should predict just the opposite, more negative transfer with low-M responses.

Greater negative transfer with high-M responses is not limited to the A-C paradigm. In fact, although the A-B, C-B (hereafter C-B) paradigm is generally regarded as a positive or, at least, zero transfer paradigm, negative transfer has been obtained with high-M responses (Jung, 1963; Twedt & Underwood, 1959). In the Jung study in which response M was varied, high-M responses led to negative while low-M responses gave positive transfer. Two sources of transfer in this paradigm can be examined here, response learning and backward associations. The lower the response M, the greater the response learning necessary in this paradigm. Whereas positive transfer is obtained when response learning is needed, negative transfer is found with high-M responses which require little or no response learning.

The backward associations of the C-B paradigm, according to Twedt and Underwood (1959), constitute an A-C paradigm in which the first-list backward associations (B-A) compete with those formed on the second list (B-C). Just as it was hypothesized earlier that response M influences first-list response availability of forward associations in the A-C paradigm, it may be argued that response M also directly increases the availability of first-list backward associations in the C-B paradigm. Thus the greater availability of interfering back-

ward associations plus the absence of response learning leads to negative transfer with high M whereas low-M responses lead to positive transfer in the C-B paradigm due to the presence of transferable response learning and lower availability of competing backward associations from the first list.

In summary, it has been argued that a consideration of the "fate" of the first-list associations during the learning of the second list of the A-C paradigm provides an associationistic explanation of the effects of first-list overlearning and response M on transfer in this paradigm. Overlearning, though eventually checked by differentiation, leads to greater negative transfer by increasing availability of competing first-list associations during second-list learning. These two factors counteract each other, thus differences in negative transfer in this paradigm as a function of degree of first-list learning are small.

Higher response M leads to greater negative transfer also by increasing availability of the first-list associations which subsequently interfere with the learning of the second list.

No need is seen for invoking cognitive structures to account for transfer effects in verbal learning. This is not to deny the possible usefulness of such concepts in explaining learning sets, discrimination-reversal learning, concept identification, and the like.

REFERENCES

- ATWATER, S. K. Proactive inhibition and associative facilitation as affected by degree of prior learning. *Journal of Experimental Psychology*, 1953, 46, 400-404.
- BARNES, JEAN M., & UNDERWOOD, B. J. "Fate" of first-list associations in transfer theory. *Journal of Experimental Psychology*, 1959, 58, 97-105.
- BESCH, N. F., & REYNOLDS, W. F. Associative interference in verbal paired-associate learning. *Journal of Experimental Psychology*, 1958, 55, 554-558.
- BIGGS, G. E. Retroactive inhibition as a function of degree of original and interpolated learning. *Journal of Experimental Psychology*, 1957, 53, 60-67.
- JUNG, J. Transfer of training as a function of degree of first-list learning. *Journal of Verbal Learning and Verbal Behavior*, 1962, 1, 197-199.
- JUNG, J. Effects of response meaningfulness (m) on transfer of training under two different paradigms. *Journal of Experimental Psychology*, 1963, 65, 377-384.
- MANDLER, G. From association to structure. *Psychological Review*, 1962, 69, 415-427.
- MANDLER, G., & HEINEMANN, S. H. Effect of overlearning of a verbal response on transfer of training. *Journal of Experimental Psychology*, 1956, 51, 39-46.
- MERIKLE, P. M., & BATTIG, W. F. Transfer of training as a function of experimental paradigm and meaningfulness. *Journal of Verbal Learning and Verbal Behavior*, 1963, 2, 485-488.
- PORTER, L. W., & DUNCAN, C. P. Negative transfer in verbal learning. *Journal of Experimental Psychology*, 1953, 46, 61-64.
- POSTMAN, L. Retention of first-list associations as a function of the conditions of transfer. *Journal of Experimental Psychology*, 1962, 64, 380-387. (a)
- POSTMAN, L. Transfer of training as a function of experimental paradigm and degree of first-list learning. *Journal of Verbal Learning and Verbal Behavior*, 1962, 1, 109-118. (b)
- POSTMAN, L., & RILEY, D. A. Degree of learning and interserial interference in retention. *University of California Publications in Psychology*, 1959, 271-396.
- SIPOLA, E. M., & ISRAEL, H. E. Habit interference as dependent upon stage of training. *American Journal of Psychology*, 1933, 45, 205-227.
- SPENCE, JANET T. Associative interference on paired-associate lists from extraexperimental learning. *Journal of Verbal Learning and Verbal Behavior*, 1963, 2, 329-338.
- THUNE, L. E., & UNDERWOOD, B. J. Retroactive inhibition as a function of degree of interpolated learning. *Journal of Experimental Psychology*, 1943, 32, 185-200.
- TWEDT, HELEN M., & UNDERWOOD, B. J. Mixed vs. unmixed lists in transfer studies. *Journal of Experimental Psychology*, 1959, 58, 111-116.
- UNDERWOOD, B. J. The effect of successive interpolations on retroactive and proactive inhibition. *Psychological Monographs*, 1945, 59(3, Whole No. 273).
- UNDERWOOD, B. J., & POSTMAN, L. Extraexperimental sources of interference in forgetting. *Psychological Review*, 1960, 67, 73-95.

(Received May 5, 1964)

SUBJECTS DO THINK:

A REPLY TO JUNG'S COMMENTS¹

GEORGE MANDLER

University of Toronto

Jung's critique is found wanting on 2 points: (a) the definition of overlearning and the effects that a structural position actually predicts, (b) the suggestion that differentiation is an associative concept. The data Jung discusses are found to be consistent with Mandler's original position, as is the notion of "differentiation," and it is suggested that Ss do think, even in paired-associate experiments.

Jung's (1965) comments raise a number of issues concerning the effects of overlearning on the A-B, A-C paradigm in verbal learning. There are a number of minor points that cannot and need not be dealt with at length, such as the notion that the subjects in a paired-associate experiment learn single responses to single stimuli, and the fact that Underwood and Postman's (1960) data indicate that for the time interval typically used in the A-B, A-C paradigm the letter-sequence effect does not predict lower availability for low-M responses.

I will restrict my present comments to two issues: What does an overlearning-structure position (Mandler, 1962) predict? What is the status of the concept of differentiation?

Overlearning and Transfer

Jung cites various studies—and in particular in relation to the A-Br paradigm and the retroactive- and proactive-inhibition (RI and PI, respectively) studies—that do not involve overlearning in my sense of the term. It should be quite clear that I consider overlearning to be a function of the number of trials following the attainment of completely correct performance. Thus, overlearning cannot be defined as a percentage function of trials to criterion as Postman (1962), for example, has done. Following this reasoning, one would obtain 200% overlearning after

six trials of acquisition if a list only takes two trials to reach criterion. It is surprising that both Jung in the present paper and other investigators have overlooked this definition of overlearning. If enough overlearning is introduced, even the A-Br condition should yield the expected results.

What predicted effects can be derived from an overlearning-structure position? I am partly at fault if there is any misunderstanding. But it should be fairly clear that overlearning of A-B should produce *less* negative transfer and probably zero transfer—with the appropriate controls—in the A-B, A-C paradigm. This point was made in my earlier paper on this problem: "The present analysis therefore predicts less interference with high degrees of learning [Mandler, 1954, p. 241]" as well as in 1962: "When . . . negative effects change to zero or positive effects . . . it is appropriate to invoke some [other than associative] hypothesis [p. 425]." The use of the term "positive" in the last quotation is possibly incorrect. All that is implied is that there should be a reduction in negative transfer following high degrees of overlearning of the A-B list. Thus, the studies by Jung (1962) and Postman (1962) both supply positive evidence for this prediction. Spence's study (1963) is, at worst, neutral in this regard. The structural position predicts that a first list of strong associates should interfere *less* with the second list than a first list of weak associates would. In fact,

¹ Preparation of this paper was supported by National Science Foundation Grant GB-810.

Spence's data show 7% and 21% negative transfer for two different lists, which *at least* in the first case is the predicted low degree of negative transfer.

One additional question needs to be clarified: What is it that is overlearned? Many different structures may be acquired, and, for example, either the specific response or a set of responses may be overlearned. In the latter case, the structure is the structure of all responses in a list; the subject can think of the list as a unitary set. This will be an important point when we return to a discussion of differentiation, but for the time being I want to apply this distinction to the problem of meaningfulness. It may be the case that, when we are dealing with previously overlearned, integrated, discrete responses—as all high-M words are—the subject is faced with a pool of highly similar integrated responses ($B + C$) which occur in the A-B and the A-C list. The individual *items* have been previously learned and are structurally available and are part of the set of high-frequency words. In the case of such high-meaningfulness lists, overlearning will produce a structural differentiation of the B list as against the C list. Thus the effect of meaningfulness will depend on what structure the subject has learned. Only when adequate overlearning has made it possible for the subject to differentiate Set B from Set C will negative transfer be reduced; otherwise intrusions among a set of meaningful words will continue to occur. With low meaningfulness or with yet-to-be-integrated responses, the subject may need only to make a decision with respect to a particular A—should he or should he not suppress the structurally available B and make (learn) the C response.

Differentiation Is Structure

Jung claims that response competition, response availability, and response differentiation are adequate to account for the decrease in negative transfer following overlearning. I have no arguments, of course, about the first two factors (cf. Mandler, 1954). And even a cursory examination of the notion of dif-

ferentiation in the two papers that have invoked it to account for the overlearning effect (Jung, 1965; Postman, 1962) reveals that it is essentially identical with my nonassociative concept of structure. From the evidence presented it is obviously not what Jung claims it to be; an associative notion.

First let me clear away some confusion in Jung's description of differentiation. He says that "as the amount of practice on one list relative to that on the other increases, differentiation may be said to occur [p. 320]." Does Jung really want to maintain that differentiation produces a U-shape function as the learning of the second list in the A-B, A-C paradigm proceeds? Obviously his statement implies that on the first trial of the A-C list differentiation between the two lists must be large; practice on A-B relative to A-C is large. As learning on the A-C list continues, differentiation will then decline and finally rise again. I do not quite know what predictions one would derive from this deduction; they might be interesting, though Jung does not make them. However, Postman (1962) implies that differentiation does not occur until overlearning on the A-B list has taken place. Note that this is what the development of structures is about. But obviously something is needed to explain the high availability of the B responses following overlearning.

In a recent experiment in our laboratory (McMillan, 1964), we found that modified free recall of overlearned B items was at about the 92% level following second-list learning, while the B recall when the first list was learned only to criterion was about 60%, and yet there were—admittedly to our dismay—practically no differences between the overlearned and the criterion groups in the mean number of trials to criterion of learning the A-C list. In fact the overlearned group was slightly—though not significantly—better. In other words, and as we expected, response availability of the B response—as measured by the modified free-recall technique—does not predict amount of negative transfer with overlearned material. In order to handle this specific

problem, Jung (1965) and Postman (1962) introduce the concept of differentiation. But how does a subject differentiate list membership? When presented a particular stimulus item, he tends to make one of two responses; the B response or the C response. If the two lists are "differentiated," he is more likely to make the C response, if they are not "differentiated," he is more likely to give the B response. But this involves a complex operation and is not an associative phenomenon in the usual sense of the term. As I understand associative phenomena such as response competition and response availability, they do not require operations on or manipulations of the materials by the subject, they do not involve cognitive transformations. To say that the subject can differentiate list membership means that he can make a choice between two possible responses, and he makes that choice on the basis of a covert operation on the structurally available responses. More simply, the subject thinks and then decides that this is the second list, and he should not be giving a first-list response. This is what I am talking about when referring to the development of cognitive structures: "We have suggested that overlearning experience with the old response and the formation of an analogic structure permits the subject to manipulate that response, to 'think about' the problem without making overt errors [Mandler, 1962, p. 425]." I see no conflict between this statement and the following quotation from Postman (1962, p. 116) about the differentiation process: "With highly integrated response terms the discrimination and rejection of errors may occur with sufficient speed to permit the substitution of the correct response during the anticipation interval." "Rejection" is not an associative mechanism, and thus neither is differentiation which, according to Postman, makes rejection possible. I am concerned with the mechanism that permits rejection to take place, and we need structural representations of some sort to perform such an operation. Overlearning of an A-B list apparently both preserves the

A-B "bond" and makes it possible for the subject to manipulate the B responses.

Let me add some further evidence from the experiment by McMillan (1964). She instructed subjects to write down the first- and second-list responses as they came to mind, i.e., in Positions 1 and 2. Thus, order of recall scores were available, and it was found that out of 8 high-frequency responses, i.e., from 8 pairs, subjects who learned the first list to criterion gave 1.8 B responses and 5.8 C responses in Position 1, while subjects who had overlearned the A-B list for 15 trials gave 5.6 B responses and 2.4 C responses in Position 1. Despite the fact that there were no significant differences in negative transfer, either in terms of trials to criterion or number of correct responses, the overlearned subjects had been able to maintain and remember the A-B "bond" at high "strength." Thus, even strongly maintained associations between A and B, associations (not just availabilities) that survive A-C learning, can be manipulated and produce no more negative transfer than relatively weak A-B "bonds" that do undergo unlearning during second-list learning.

Finally, I would like to refer Jung to the literature on the effect of anticipation intervals in the A-B, A-C paradigm. Nodine (1963) and others have shown that negative transfer decreases as a function of the anticipation interval employed. If one accepts the notion that in the A-B, A-C paradigm the subject must *suppress* the B response during second-list learning, then an increase in the length of the anticipation interval would reduce the time pressure and make suppression of the inappropriate first-list response easier. Similarly, first-list overlearning makes such suppression easier. In other words, either long anticipation intervals where the subject can think about the appropriate response, or overlearning, where the subject has the inappropriate response readily available and readily suppressible, produce less negative transfer.

Postman (1962) and, by implication, Jung invoke "differentiation" to explain how a subject rejects (or suppresses) an incorrect response. That is the central

problem, and it cannot be solved by introducing terms that sound better than others in order to salvage the consistency of a particular theoretical school. I think we should use concepts of whatever parentage and not be particularly concerned with such salvage. The use of the term "structure" suggests that the subject is able to manipulate several responses available at the same time. The study of human verbal learning is part of the problem of thinking, and we should be concerned how subjects learn to manipulate covert responses, how they identify them, reject them, avoid interference, how—in short—thinking proceeds.

I believe that the invocation of "differentiation" is additional support for a nonassociative explanation of the effect of overlearning on the A-B, A-C paradigm. Jung is quite right when he points out that reduction in the negative transfer is not an easily found phenomenon. But true overlearning is not a frequently used condition in the typical paired-associate experiment. The conjunction of overlearning and cognitive structures was introduced because overlearning appeared to be one of the conditions—occurring in the natural habitat of the human organism—that leads to structural mechanisms. To believe that man thinks only when he is out of the laboratory and that we should ignore this phenomenon as much as possible inside the laboratory seems to be

foolhardy. Subjects do think, even when learning paired-associate lists.

REFERENCES

- JUNG, J. Transfer of training as a function of degree of first-list learning. *Journal of Verbal Learning and Verbal Behavior*, 1962, 1, 197-199.
- JUNG, J. Comments on Mandler's "From association to structure." *Psychological Review*, 1965, 72, 318-322.
- McMILLAN, LINDA R. The role of response suppression and overlearning in associative interference. Unpublished master's thesis, University of Toronto, 1964.
- MANDLER, G. Response factors in human learning. *Psychological Review*, 1954, 61, 235-244.
- MANDLER, G. From association to structure. *Psychological Review*, 1962, 69, 415-427.
- NODINE, C. F. Stimulus durations and stimulus characteristics in paired-associates learning. *Journal of Experimental Psychology*, 1963, 66, 100-106.
- POSTMAN, L. Transfer of training as a function of experimental paradigm and degree of first-list learning. *Journal of Verbal Learning and Verbal Behavior*, 1962, 1, 109-118.
- SPENCE, JANET T. Associative interference on paired-associate lists from extraexperimental learning. *Journal of Verbal Learning and Verbal Behavior*, 1963, 2, 329-338.
- UNDERWOOD, B. J., & POSTMAN, L. Extra-experimental sources of interference in forgetting. *Psychological Review*, 1960, 67, 73-95.

(Received June 8, 1964)

PSYCHOLOGICAL REVIEW

TRANSFER OF VERBAL PAIRED ASSOCIATES¹

EDWIN MARTIN²

University of Iowa

From the current literature it is possible to identify 2 processes underlying the acquisition of verbal paired associates: response learning and association formation. It is also apparent from the literature that a complete treatment of association formation must take into account association directionality. Altogether, then, 3 "things" are seen to evolve during the acquisition of a paired-associate list: response availability, forward associations, and backward associations. The thesis of the present research is that what is transferred from the 1st to the 2nd task in a paired-associate transfer situation is some combination of these 3 effects. Utilizing the coordinate system invented by Osgood in which all the transfer paradigms can be arranged, 3 transfer surfaces are proposed which describe how each of the 3 effects is transferred individually. Applications to extant transfer literature are made. It is found that the principal results of nearly all experiments utilizing the A-B, C-D control paradigm can be accounted for.

The problem of transfer of verbal paired associates originated with Müller and Schumann (1894), whose law of associative inhibition was essentially a specification of the negative transfer paradigm A-B, A-D, and Müller and Pilzecker (1900), who de-

veloped the method of "right associates." A closely related problem, one with which transfer is, in many respects, inextricably intertwined, is that of retroactive inhibition, which also originated with Müller and Pilzecker (1900) in conjunction with their perseveration theory of reproduction inhibition. In 1949, Osgood attempted to organize the many facts and insights that had been accruing in this general problem area since Müller's time by proposing a transfer and retroaction surface. Thus the period from Müller to Osgood is a historical package, so to speak, with the most articulate summary being Osgood's paper. The purpose of the present paper is a reorganization based on additional facts and insights, some of which have accumulated since 1949, with the restric-

¹ This article derives in part from a doctoral dissertation submitted to Graduate College, University of Iowa. The author gratefully acknowledges the advice and guidance of R. W. Schulz. The valuable comments of A. W. Melton are also recognized. Preparation of the manuscript was supported by the Advanced Research Projects Agency, Department of Defense, monitored by the Air Force Office of Scientific Research, under Contract No. AF 49(638)1235 with the Human Performance Center, University of Michigan.

² Now at Human Performance Center and Department of Psychology, University of Michigan.

tion that only transfer problems will be considered.

In an analysis of the effects of having learned one set of paired associates on the subsequent learning of another, four experimentally controllable variables emerge for which there are sufficient data to draw definitive conclusions. Two of these, interlist stimulus and interlist response similarity, have received considerable attention and were incorporated into Osgood's (1949) transfer theory. The other two, degree of first-task learning (L_1) and response meaningfulness (M), *have been treated experimentally but not explained theoretically*. To grasp the empirical relationships among these four variables, it seems best to first examine Osgood's theory and then to describe how L_1 and M underlie systematic departures.

Osgood's (1949) contribution to transfer theory can be seen to resolve into two components: (a) the invention of a coordinate system in which all paired-associate paradigms can be arranged, and (b) a summary of available data in the form of a surface which describes how amount of transfer and position in the coordinate system are related.³ The coordinate system has

³ At approximately the same time, and apparently independently, Underwood (1949) made an equivalent proposal based on the same data, plus expectations from generalization theory, in the form of four curves. His curves are essentially the surface edges given by Osgood.

the important feature of formally distinguishing the separate roles of interlist stimulus and interlist response dissimilarity. These two variables are expressed as orthogonal axes, the X_S and X_R axes of Figure 1F, with the origin representing complete similarity, or identity. Thus, for example, the negative transfer paradigm A-B, A-D has the coordinates $X_S = 0$ and X_R some sufficiently large value to indicate complete dissimilarity (or unrelatedness) between the responses of the first- and transfer-task lists. All transfer paradigms are assigned a unique *position in joint accordance with the dissimilarity between the stimuli and the dissimilarity between the responses of the two lists*. Surface points above the X_S - X_R plane represent positive transfer; points below, negative transfer.

A matter to be taken into account before proceeding further is that of the so-called "opposed" relationship between the two tasks. With stimuli identical, for example, the responses of the first and transfer tasks can be identical (A-B, A-B paradigm), similar (A-B, A-B'), neutral (A-B, A-D), or opposed. A number of investigators, including Osgood, have used this last condition. As an example of the opposition relationship, *sickly* is similar to *pale*, but *healthy* is opposed. Usually, response opposition produces less negative transfer than the neutrality relationship of the A-B, A-D

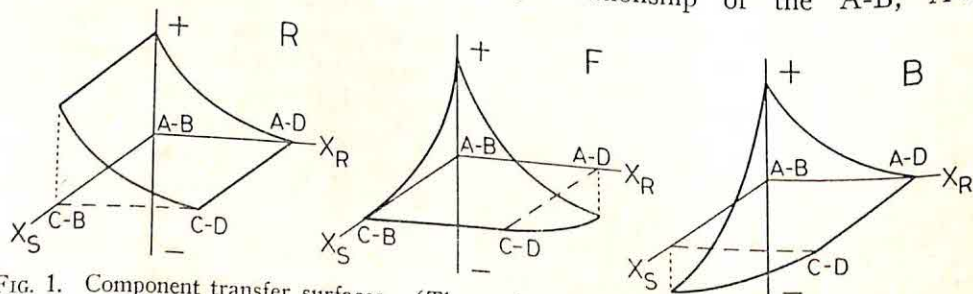


FIG. 1. Component transfer surfaces. (The surfaces R, F, and B represent the transfer of response availability, forward associations, and backward associations, respectively.)

paradigm (e.g., Bugelski & Cadwallader, 1956; Wimer, 1964).

If two items are judged opposite in meaning, they must perforce be related; otherwise, opposition could not be established. Generally speaking, the axes of Osgood's coordinate system are seen as representing similarity in meaning. However, it would be more consonant with the present status of interference theory to suppose that the axes represent associative relatedness. Similarity would clearly be included; but, more importantly, opposition would thereby be placed at points *on the axes between identity and complete dissimilarity instead of beyond the complete dissimilarity point.* To so recognize opposition as a category of relatedness achieves two reasonable objectives: (a) It places data points arising from opposition relationships (between Tasks 1 and 2) in an area of the coordinate system where transfer is known to be less than maximally negative. In the experiment by Bugelski and Cadwallader (1956), the response-opposition condition would thus interpose between the identical (A-B, A-B) and similar (A-B, A-B') conditions, but closer to the similar condition; in the Wimer (1964) experiment, the response-opposition condition would fall between the similar (A-B, A-B') and unrelated (A-B, A-D) conditions. (b) It would more completely link interference interpretations of verbal learning to the transfer situation by removing the restriction that the only relationship represented by the coordinate axes is that of strict similarity. Associative relatedness is the basis for both synonymy and antonymy. There can thus be no paradigm on the dissimilarity axes beyond the completely unrelated paradigms (A-B, C-B for stimuli; A-B, A-D for responses). This is somewhat of a departure from Osgood's

original conceptualization, but a departure which nevertheless seems necessary if thinking about what is transferred is to square with current thought on what is acquired when learning takes place.

At positions where the surface is either above or below the X_S - X_R plane, positive or negative transfer, respectively, is supposed to obtain. A perusal of the transfer studies whose designs constitute a test of Osgood's surface, however, reveals that although the general shape of the surface is fairly well substantiated, there are a number of systematic discrepancies which can *be attributed to specific experimental factors, namely, L_1 and response M.* (In the ensuing discussion, only experiments utilizing the A-B, C-D control paradigm for nonspecific effects will be cited; thus amount and sign of transfer at other positions in the coordinate system are always with respect to transfer-task performance at the A-B, C-D position.)

The effect on transfer-task performance of increasing L_1 is either to decrease positive transfer or to increase negative transfer, depending upon the position in the coordinate system. Bruce (1933) found that as L_1 increased from 6 trials to 12, the amount of positive transfer in the A-B, C-B paradigm decreased, while in the A-B, A-D paradigm the amount of negative transfer increased. In Jung's (1962) experiment, transfer progressed from positive to negative in the A-B, C-B paradigm and became more negative in the A-B, A-D paradigm as L_1 increased from a 3/6 to a 6/6 + 5 criterion. Dean and Kausler (1964) have shown a striking shift from positive transfer with a first-task criterion of $L_1 = 3/6$ to negative transfer with $L_1 = 6/6 + 5$ in the A-B, C-B paradigm. Further substantiation of this role of degree of L_1 is provided by

Spiker (1960) for the A-B, A-D paradigm and by Postman (1962b) for several paradigms.

If overlearning of the first task is considered, the foregoing clear picture of the role of degree of L_1 becomes somewhat clouded. Postman (1962b), while reporting an increase in negative transfer with a change from a 6/10 to a 10/10 criterion for L_1 in both the A-B, C-B and A-B, A-D paradigms, finds a sharp reduction in amount of negative transfer for a 10/10 + 50% criterion. Mandler and Heinemann (1956) report negative transfer in the A-B, C-B paradigm after 10 errorless L_1 trials but positive transfer after 30 and 50 such trials, with a return to negative transfer after 100 trials. In the A-B, A-D paradigm, transfer remains negative throughout but shows fluctuation similar to that in the A-B, C-B. These results are difficult to reconcile both with each other and with data not involving overlearning. Mandler (1962) has presented a systematic argument for negative transfer as a V-shaped function of degree of L_1 , transfer becoming more negative then less negative in the A-B, A-D paradigm. It should be noted, however, that in making that argument, Mandler cites his own data (Mandler & Heinemann, 1956) but leaves out the data point for his greatest-degree-of-overlearning condition, a data point which denies the V-shaped function. Other data cited by Mandler in support of a V-shaped function are not relevant in the present context because they do not include A-B, C-D control paradigms for nonspecific effects. Thus, although degree of L_1 inversely affects transfer in the A-B, A-D and A-B, C-B paradigms in a regular manner at ordinary degrees of L_1 , a definitive conclusion regarding the role of overlearning does not seem possible at this point.

Where M is taken as a generic term referring to the core attribute of verbal units reflected by such highly inter-correlated scales as meaningfulness, association value, and pronunciability (Noble, 1963), the effect of first-list response M on transfer-task performance is to increasingly impede that performance as M is increased. Jung (1963) showed that when responses are low- M trigrams, considerable positive transfer and barely negative transfer are observed in the A-B, C-B and A-B, A-D paradigms respectively; but when responses are high- M trigrams, transfer in the A-B, C-B becomes nearly zero, and transfer in the A-B, A-D becomes very negative. Varying M over three levels (low—CCCs, medium—CVCs, high—words), Merikle and Battig (1963) showed that as M increases, transfer in the A-B, A-D paradigm decreases from barely negative to considerably negative.

Further evidence on this matter can be had by categorizing available data according to whether transfer is positive or negative in the A-B, C-B paradigm. If this is done, it turns out that all studies yielding positive transfer involve trigrams as responses (Bruce, 1933; Jung, 1963; Mandler & Heinemann, 1956), whereas those yielding negative transfer involve meaningful responses, such as adjectives (Dallet, 1962; Harcum, 1953; Kausler & Kanoti, 1963; Keppel & Underwood, 1962; Twedt & Underwood, 1959; Wimer, 1964). An exception to this categorization is the Bugelski and Cadwallader (1956) study, which reports positive transfer at the A-B, C-B position with words as response materials. Their procedure involved dropping out pairs from the first-task list once given correctly twice in succession. Further, they used visual patterns as stimuli.

Thus degree of L_1 and response M

appear to be major experimental factors responsible for systematic deviations from Osgood's surface. There are, of course, other factors whose effects will eventually have to be accounted for, for example, intralist similarity, list length, and time between L_1 and the transfer task. At the present time, little is known about the effects of the first two on the transfer task; and, adhering to the restriction of considering only transfer data which includes the A-B, C-D control, only one study, Newton and Wickens (1956), is known to the present writer concerning the last. If one is willing to compare percent saving (the percent the difference in the number of trials to criterion for the two tasks is of the number for the first task) in the A-B, A-D paradigm (their Experiment I) with percentage of saving in the A-B, C-D paradigm (their Experiment II), it turns out that the differences obtained by subtracting the latter from the former for intertask delays of 0, 24, and 48 hours is -26, -8, and +5. Thus for the A-B, A-D paradigm there is evidence that the associations which are to interfere during transfer-task learning are lost over time. The studies customarily cited on this phenomenon, however, either do not employ the A-B, C-D control for nonspecific effects (Bunch & McCraven, 1938) or deal only with the nonspecific effects of the A-B, C-D paradigm (Hamilton, 1950).

Essentially, while the coordinate system is definitional in nature, the surface proposed by Osgood is an empirical law, an induction from the facts. In view of the growing complexity of the facts as revealed, at least, by the roles of L_1 and M , however, a revised induction does not seem feasible; instead it would appear easier to turn to the implications embedded in already extant verbal learning theory for an

improved formalization of transfer phenomena. On this view, two developments are of special interest: the two-phase conceptualization of verbal learning and the notion of association bidirectionality.

The idea that verbal learning may profitably be viewed as more than one process seems to have originated with Thorndike (1932), who notes that "other things being equal, connections are easy to form in proportion as the response is available [p. 343]" and that "much of learning consists in making certain responses more available [p. 347]." Current expression of this notion is well exemplified by Underwood and Schulz (1960): "Logically speaking, the acquisition of a serial or paired-associate list can be divided into two stages. The first will be called *response-learning* or *response-recall* stage. It occurs temporally prior to the second stage which will be called the *associative* or *hook-up stage* [p. 92]."

The utility of distinguishing two phases of the total verbal learning process is readily discernible in the literature. For example, Underwood, Runquist, and Schulz (1959), Horowitz (1961), and Carterette (1963) have demonstrated that the two phases can be discriminated experimentally on the grounds that a single variable differentially affects certain response measures separately identifiable with the two phases. It was found that while variations in intralist response similarity had a direct effect on response learning (as measured by recallability), association formation (as measured either by responding to specific stimulation or by ordering performance) was inversely affected.⁴

⁴ Caution must be exercised in claiming that high intralist response similarity facilitates response learning. When similarity among items is produced by means of either

A more detailed consideration of these two components of learning, plus others, has been given by McGuire (1961): He argues cogently for their utility and demonstrates experimentally their distinguishability.

Since it is meaningful to view the acquisition of verbal paired associates as a composite of two processes, response learning and association formation, it makes sense to conclude that at the end of L_1 in a transfer situation there are at least two "things" available for transfer to the second, or transfer, task; namely, response availability and associative connections, the former being the product of first-task response learning, the latter of first-task association formation.

Admitting that associative connections are available for transfer from task to task entails a consideration of the rather extensive literature on backward associations. Cognizance of this problem apparently begins with Ebbinghaus (1885), who wrote that "as a result of the learning of a series cer-

tain connections of the members are therefore actually formed in a reverse as well as in a forward direction [p. 112]." Most of the early interest in backward associations was centered on the problem of remote associations in serial learning, the first application to transfer apparently not occurring until the appearance of Harcum's paper in 1953. Harcum's finding that backward associations are transferred has since been confirmed by Murdock (1958) with different materials, with further evidence provided by Keppel and Underwood (1962); hence a generality seems to be in order. The conclusion, then, is that backward associative connections are a component of what is transferred in a transfer situation and must be included in a formalization of transfer phenomena.

A COMPONENT-SURFACES CONCEPTUALIZATION

It has been argued that there are three components to that which is transferred from one paired-associate task to another, namely, response availability, forward associations, and backward associations. An immediate implication of this position is that all of the ordinary measures of transfer (number of trials to criterion on the transfer task; number of correct responses, or errors, in so many transfer-task trials; etc.) reflect what is best called a net effect. In other words, a single transfer surface whose ordinate values are given by any one of the standard measures is essentially a net transfer surface and presumably can be analyzed into other, more specific component surfaces. This means that, since there are three effects transferred, there should be three transfer surfaces, each describing how its particular effect is transferred as a function of position in the coordinate system. The specification of such a triad of surfaces,

(a) generating a list of trigrams from a limited alphabet or (b) basing similarity on the notion of synonymity, the possibility exists that given one response some of the others can be inferred. In such situations, it becomes problematic whether observed facilitation is effected by current learning or by transfer of previously acquired rules or strategies.

It is not reasonable to suppose that response learning is the sole source of the response availability measured by recall. Strictly speaking, response learning is the serial acquisition of the elements of the response (and hence is itself an association-formation process). Response availability, however, involves also the contextual associations that support recall. Thus the recall measure is not a pure measure of response learning. In the transfer situation, context associations which tend to elicit responses are indistinguishable from specific transferred associations, hence in future references to response availability what is intended is availability arising from response learning only.

at least schematically, follows from general paired-associate interference theory.

Consider first how response availability (that due to response learning) transfers from one paired-associate task to another. Any response learning evolved during the first task must transfer maximally to the transfer task of paradigms at positions in the coordinate system where first- and transfer-task responses are identical. This transfer is not perfect to the extent that response availability includes context and/or other nonparadigmatic associations not carrying over; in fact, such associations may even be a source of interference. Response availability due to first-task response learning, however, has a maximum and equal positive effect along the X_S axis, that is, at all positions from A-B, A-B to A-B, C-B. At the other extreme, none of the response learning of the first task is appropriate to the transfer task of paradigms involving completely new transfer-task responses. Response availability acquired extraexperimentally may certainly apply but not response availability from first-task learning. Therefore, a surface representing the transfer of first-task response learning must coincide with the X_S - X_R zero-transfer plane at positions from A-B, A-D to A-B, C-D.

Regarding positions between the maximally positive-transfer loci from A-B, A-B to A-B, C-B and the zero-transfer loci from A-B, A-D to A-B, C-D, intermediate amounts of response availability obtain. These varying degrees of transferred response availability are seen as arising from the applicability of first-task response learning to whatever characteristics or components of the transfer-task responses are responsible for the similarity (relatedness, including opposition) between the first- and transfer-task

responses. This argument can be summarized by a hypothetical surface, called the R surface for response availability, and is shown in Figure 1R.

The second component of paired-associate acquisition, forward association, requires some degree of similarity between first- and transfer-task stimuli in order to induce a nonzero effect. If associations A-B are acquired in the first task, then to the extent to which the stimuli of the transfer task are similar to A, the associates of A from the first task are elicited in transfer-task learning. Thus, those positions on a line between A-B, C-B and A-B, C-D cannot involve transfer of forward associations because completely new stimuli characterize the transfer task. In progressing from the A-B, C-B position toward the A-B, A-B position in the coordinate system, first-task forward associations become increasingly more useful in the transfer task, until the A-B, A-B position is reached where the transfer task is but a continuation of the first task.

As the A-B, A-D position is approached (from any direction), however, the forward associations acquired in the first task become increasingly stronger sources of interference in transfer-task learning. This interference is inferred from the fact that acquisition of A-D associations involves the extinction of A-B associations (Barnes & Underwood, 1959; Briggs, 1954; Postman, 1961, 1962a). As will be pointed out below, the literature is unanimous that relative to the A-B, C-D control paradigm the A-B, A-D yields negative transfer. The point to be made here is that this negativity cannot be due to either transferred response availability or transferred backward associations (because the responses of the transfer task are completely new) and hence must be due to interfering forward associations. The

F surface, depicting the transfer of forward associations, is shown in Figure 1F. It is essentially the surface proposed by Osgood (1949).

The third, and final, component to be considered is that of backward association. The surface to represent this effect must necessarily be symmetric in general form, if not in magnitude, with the F surface where the axis of symmetry is the line between the A-B, A-B and A-B, C-D positions. Arguing in the reverse direction, there must be some degree of similarity (relatedness, including opposition) between the responses of the first and transfer tasks in order for backward associations acquired in the first task to be elicited in the transfer task. Now it has already been pointed out that backward associations are a component of what is acquired in paired-associate learning; hence if, as in the A-B, C-B paradigm, the backward B-A associations of the first task are elicited during transfer-task learning, then the acquisition of the backward B-C associations will be impeded. Thus the surface representing the transfer of backward associations must be negative at the A-B, C-B position. At the A-B, A-B position, however, the backward associations to be acquired in the transfer task are the same as those already practiced in the first task, hence the surface there must be positive. On a line between A-B, A-D and A-B, C-D, the responses are completely different in the transfer task, that is, the stimuli capable of eliciting the backward associations of the first task are absent, hence the surface must coincide with the X_S - X_R zero-transfer plane. Such a surface is shown in Figure 1B.

The three surfaces of Figure 1 are not entirely empirical (as was Osgood's single surface) because the three concepts represented by them are indeed primarily concepts. The surfaces are

intended to be representative of expectations as to the nature of the components of the two-task paired-associate situation (transfer) arising from the general interference theory of the single-task paired-associate situation.

The locus of action of the experimental variables L_1 and M can now be seen somewhat more clearly. In the degenerate case where $L_1 = 0$, that is, where no first-task learning obtains, all surfaces collapse into the X_S - X_R zero-transfer plane: Nothing is transferred. With successive first-task trials, the associations which are to transfer become stronger, and the F and B surfaces assume their characteristic curvature. For example, at the A-B, C-B position, as L_1 increases, the negativeness of the B surface increases, presumably asymptotically, until a maximum is reached (see earlier discussion of overlearning). With respect to the R surface, L_1 trials serve also to develop its curvature. In situations where first-task responses are of the highest M (e.g., familiar nouns), only a few trials are required for full development of the R surface. Thus L_1 and M enter the analysis as factors determining the developmental (over trials) dispositions of the surfaces *relative to the X_S - X_R zero-transfer plane*. These relationships will emerge more clearly when applications of the model are discussed below.

Each surface is thus a different function of the same two variables, X_S and X_R . For example, the R surface would be described by some function, f , as

$$R = f(X_S, X_R; M, L_1),$$

where M and L_1 are parameters representing the response- M and degree-of- L_1 characteristics, respectively, of the transfer situation. If response M , say, were to be studied functionally at some fixed position in the coordinate

system, then resulting variations in response availability would be represented by

$$R = f(M; X_S, X_R, L_1),$$

where f is the same function as before, but M is now taken as the variable and X_S , X_R , and L_1 as the parameters. Thus there are multiple experimental approaches to the ultimate specification of the three functions relating the component surfaces to X_S and X_R . The most formidable obstacle between the present state of transfer theory and such a goal, however, is the unavailability of a suitable similarity (relatedness) measure for the X_S and X_R dimensions. If such a measure cannot be evolved, then the Osgood coordinate-system view of transfer phenomena cannot develop into a quantitative theory.⁵

The decomposition of transfer phenomena into component effects has been given brief consideration by Postman (1962b). He distinguishes four sources of transfer: (a) learning to learn and warm-up, or nonspecific effects; (b) response learning, which is represented by the R surface in the present formulation; (c) associative interference, which has here been formally subdivided into the F and B component surfaces; and (d) differentiation between lists, a characteristic of transfer tasks represented by the X_S and X_R values assigned to the paradigm involved. Regarding the last source, if X_S and X_R are to adequately represent

differentiation between lists, they must necessarily include more than scaled dissimilarity. Undoubtedly, Postman means to include the type of differentiation characterized by increased certainty as to which of the two tasks a given response item belongs. Postman draws further attention to the fact that second-task learning also has two phases and that the several effects being transferred from the first task should be expected to differentially affect them.

Before proceeding to applications of the component-surfaces conceptualization to specific transfer situations, it is necessary to take one more step on the theoretical level: In addition to the defined coordinate system and the assumed component surfaces, there needs to be a statement about the temporal relationship between the two phases of learning. The role of such a statement would be to provide a conceptual mechanism by means of which the parameters of the transfer problem can specify dispositions of the component surfaces *relative to each other*. For example, suppose it were assumed that the response-learning phase precedes the association-formation phase; then for a sufficiently low degree of L_1 very little association formation could occur. Therefore, at a position such as the A-B, C-B where the sign of transfer depends upon the net effect of the positive R and the negative B effects (see Figure 1), positive transfer would be predicted because no B effects could have developed. But this is an example of a successful application of the complete component-surfaces conceptualization to actual data and hence is putting the cart before the horse, so to speak. If the original plan of attack is to be adhered to, a statement about the temporal relationship between the two phases should be evolved outside of transfer theory.

⁵ It is of interest to note that in a component model for stimulus compounding and generalization presented by Atkinson and Estes (1963) a statement can be derived which says that in general the similarity relation between two verbal units is not symmetric; for example, abc may be more similar to ab than ab is to abc . If this is in fact the case, it would make a difference which list goes into which task in a transfer problem, hence introducing directionality into dissimilarity measure.

The facts insist that there is no clear temporal distinction between the response-learning and association-formation phases in verbal acquisition. Three studies of particular interest on this score are those of Peterson and Peterson (1959, Experiment II), Crothers (1962), and McGuire (1961). Peterson and Peterson studied the short-term retention of consonant syllables and discovered that the dependent probabilities of subsequent letters being correct, given that the preceding letters were correct, increased with the number of repetitions before recall. In other words, parts of a unit may become available before the unit is available as a whole. Crothers, working in a paired-associate situation, found that when response items comprise more than one component, association formation may begin between the stimulus and whatever components are available before the entire response is available as a unit. This conclusion for the paired-associate situation finds further exemplification in the work of McGuire. The import of these studies is: Although response learning and association formation may be effectively distinguished, they nevertheless overlap; for association formation may begin as soon as at least one of the components of the response item is available.

Certainly the entire response unit must be available before a correct response can be accredited, and certainly association formation continues after responding has become correct; but these two arguments say nothing about when, relative to each other, the bulks of the two processes occur. An adequate analysis of this problem requires first an agreement as to what is meant by "the bulks of." Pending such a step, it might be noted in the Underwood, Runquist, and Schulz (1959) data that whereas the effects of intralist

response similarity on recall of the response items disappear after the first half-dozen trials, the effects on overall learning carry on to the end, an indication that response learning may be a more rapid process than association formation in some situations. For theoretical purposes, it will be assumed that, in general, response learning precedes, but overlaps, association formulation. Vindication of this step will emerge as applications are considered.

The coordinate system, the component surfaces, and the assumption just made comprise a clearly asymmetric formulation: No mention has been made of stimulus learning or integration. In view of the limited role of experimentally manipulated stimulus availability in subsequent paired-associate learning (e.g., Schulz & Martin, 1964; Schulz & Tucker, 1962), and because of the small effect of stimulus M relative to response M (e.g., Underwood & Schulz, 1960), it does not seem profitable at the present time to develop an accommodating structure within transfer theory. That certain stimulus factors will prove to be of considerable importance in paired-associate learning is clear from the work of McGuire (1961) and Shepard (1963); however, the implications for transfer situations are not yet clear.

APPLICATIONS

In order to see how the foregoing proposed conceptualization applies to extant data and generates expectations, several of the more thoroughly studied paradigms will be considered separately.

A-B, C-D. First, it is clear that the analysis of the net surface into component surfaces leaves the A-B, C-D position as the appropriate control for all other positions; in fact, its naturalness as a control is emphasized. Since the transfer task involves new responses, none of the backward asso-

ciations formed in L_1 is elicited; nor is there any "head start" on transfer-task response learning since the response familiarization effected in L_1 is completely irrelevant. Similarly, since the transfer task involves new stimuli, none of the forward associations formed in L_1 is elicited. On the other hand, the transfer task does benefit from the nonspecific transfer effects common to all the paradigms. Certainly the measure of transfer used, and hence the choice of a control, depends upon the intent of the investigator; but if the specific effects of response availability and forward and backward associations are being studied, use of the A-B, C-D control seems imperative.

A complication alluded to by Postman (1962b) regarding the adequacy of the A-B, C-D position as a control when low-M material is used should be mentioned. Although not so stated by him, the cause for concern arises when the items are, say, consonant trigrams (CCCs), thus making it very difficult to satisfy the stipulation (required for the adequacy of the A-B, C-D position as a control) that first- and transfer-task stimuli and responses are completely dissimilar. Due to the limited number of letters (21) from which CCCs can be formed, at least some formal similarity between lists of any length will obtain.

A-B, C-B. Of the three effects, R, F, and B, available for transfer at the end of L_1 , only R and B are actually transferred at the A-B, C-B position. Since the transfer task involves new stimuli, forward associations are not elicited, hence leaving in opposition the positive R and negative B effects. This much is based simply on the assumed component surfaces. In order to form an expectation as to the circumstances in which the net R-B effect will be positive or negative, it is neces-

sary to evoke the assumption that response learning precedes association formation. Thus when degree of L_1 is sufficiently low, only response learning will have occurred, to the exclusion of association formation, thereby allowing only R effects to evolve. Increasing the degree of L_1 permits the opposing negative B effects to develop. As degree of L_1 goes from low to high, the net R-B effect shifts from positive to negative. In this way, the theory can account for the observed variations in transfer at the A-B, C-B position as a function of degree of L_1 .

A similar use of the R and B component surfaces and the assumption regarding the temporal order of response learning and association formation leads to an explanation of the role of response M in transfer at the A-B, C-B position. When low-M first-task responses are used for a given degree of L_1 , L_1 is primarily taken up with response learning and only secondarily with association formation; but when high-M responses are used, such an extensive response-learning phase is not necessary, allowing more emphasis on the formation of the backward associations which will cause interference in the subsequent transfer task. An increase in response M, then, means an increased opportunity for the development of negative B effects, hence either a decrease in positive transfer or an increase in negative transfer, depending on the given degree of L_1 .

The foregoing theoretical treatment of the effects of degree of L_1 and level of response M at the A-B, C-B position is intended as an explanation of the data cited earlier as discrepant from Osgood's single surface. Several predictions for situations in which there are as yet no data can also be generated. One of these has to do with the effect of variations in first-task intralist response similarity. Since it is well

known that high intralist response similarity tends to impede overall paired-associate learning but to facilitate response learning (e.g., Feldman & Underwood, 1957; Underwood, Runquist, & Schulz, 1959), decreasing amounts of negative transfer should be associated with increasing intralist response similarity. In comparing two groups, one with high and one with low first-task intralist response similarity, during L_1 the former will have evolved more response availability (stronger R effects) than the latter in equal numbers of trials. This is because high similarity facilitates response learning. (In addition, the difference between these two groups should be larger for low-M response materials than for high.) Further, in the group with high intralist response similarity, acquisition of the associative connections (B effects) which will subsequently produce interference in the transfer task will be impeded to a greater extent than in the group with low similarity. This is because response similarity impedes association formation. Thus the high similarity group should not only have stronger positive R effects, but also less developed negative B effects, and therefore should show positive transfer relative to a low-similarity group.

Other predictions attach to the several ways of varying response M or variables relating to response M. For example, one should be able to independently manipulate the R surface by such expedients as varying the number of "chunks" (Melton, 1963) in the first-task response members or administering varying amounts of response familiarization prior to beginning the first task. The consequences of such procedures would be differential emphases on association formation during L_1 and hence variations in the net R-B effect.

A-B, A-D. At the A-B, A-D position, one sees from Figure 1 that neither R effects nor B effects are transferred, leaving only the interfering forward associations acquired during L_1 . This unqualified expectation of negative transfer is unanimously borne out by studies using the A-B, C-D control. In order to explain the roles of degree of L_1 and level of response M at this position, it is again necessary to evoke the assumption that response learning precedes association formation.

Variations in degree of L_1 , regardless of position in the coordinate system, entail variations in degree of association formation: Since response learning precedes association formation, the time allotted to the formation of the forward associations which subsequently cause interference during transfer-task acquisition depends directly upon the amount of L_1 in excess of that required for the response-learning phase. Thus as degree of L_1 increases, the amount of negative transfer at the A-B, A-D position should increase.

A similar argument leads to the expectation of increased negative transfer with increased M of first-task responses. Although the R surface is inoperative at the A-B, A-D position, variations in response M differentially delay (inversely according to level of M) the association-formation phase of L_1 .

The foregoing derivations regarding degree of L_1 and level of response M are in complete accord with the data cited earlier.

A result which can be accounted for in terms of component surfaces but not by the single transfer surface is that of the Young and Underwood (1954) response-predifferentiation study. They studied three paradigms situated along the edge of the surface from A-B,

A-D to A-B, A-B. Half of the subjects were given predifferentiation training on the response items of their first- and transfer-task lists by means of a verbal discrimination task, the other half receiving irrelevant predifferentiation training. Those in the latter category yielded the expected gradient of decreasing positive transfer with decreasing interlist response similarity; those receiving relevant response predifferentiation yielded a markedly modified gradient, best represented by a horizontal line. In other words, relevant response predifferentiation eliminated the gradient. Now response predifferentiation can be seen to have two effects in such a situation. First, mere repetition (during the verbal discrimination task) of the items which are subsequently to be responses raises the R surface, that is, increases response availability, with the maximum effect occurring at the A-B, A-D position. Second, the X_R value for the paradigm involved is increased; this is because predifferentiation serves to increase the distinction between the responses of the two lists. The net effect of these two factors should be negative for paradigms close to the A-B, A-B position since changes in the R surface are less effective there (R is at its maximum at A-B, A-B), but the fast dropping F surface would bring about a considerable negative shift with the increase in X_R . On the other hand, the net effect on paradigms near the A-B, A-D paradigm should be positive: Paradigms near there receive nearly the full effect of increased response availability but only a negligible negative effect due to increased X_R .

Table 1 summarizes the signs of the three components of the net transfer effect for the three paradigms discussed, plus the A-B, A-B. Note that Table 1 is a tabular restatement of Figure 1 for the four positions listed.

TABLE 1
DIRECTION OF COMPONENT
TRANSFER EFFECTS

Paradigm	Component effect		
	R (Response availability)	F (Forward associations)	B (Backward associations)
A-B, A-B	+	+	+
A-B, C-B	+	0	-
A-B, A-D	0	-	0
A-B, C-D	0	0	0

*A-B, A-Br.*⁶ Although the Osgood coordinate system is entirely adequate for providing a position for every paired-associate transfer paradigm, an examination of the A-B, A-Br paradigm leads to the conclusion that the component-surfaces formulation proposed in this paper cannot be applied to a certain class of paradigms of which the A-B, A-Br is a member. Three peculiarities of this paradigm seem of importance. First, because the exact same stimuli and responses appear in both tasks, it must be that, for this paradigm, $X_S = X_R = 0$, and hence that the A-B, A-Br is assigned the same position as the A-B, A-B. But the latter represents a maximum positive-transfer situation while the former typically involves considerable negative transfer. This is because in the A-B, A-Br the transfer-list re-pairings require transferred F and B effects to be negative. Therefore, in spite of the fact that the A-B, A-B and A-B, A-Br paradigms occupy the same position in the coordinate system, the same triad

⁶ This paradigm designates transfer situations where the stimulus and response members of the transfer task are identical with those of the first except that they are re-paired. The notation was introduced by Porter and Duncan (1953), the first to study this paradigm using verbal materials. Gagné, Baker, and Foster (1950) discuss a similar situation in motor transfer problems under the name of complete reversal learning.

of component surfaces cannot describe the transfer effects of both.

Second, although each transfer-list stimulus has a new response, the responses are not new in the sense that they are in the A-B, A-D paradigm; and further, the old responses are still present somewhere in the list, not completely removed as in the A-B, A-D paradigm. To understand the import of such an arrangement, consider a single transfer-task pair where the stimulus member is the same as it was in the first task but where the response member is different. There are two kinds of response items which may be assigned: a completely *new* response item, that is, an item not used in the first task (1); or an *old* response item, one already used with a different stimulus in the first task (2). But both of these kinds of response items may be assigned to the stimulus under either of the following two conditions: (a) where the original response item, the one that went with this stimulus in the first task, is completely *removed* from the transfer list; or (b) where the original response item is *retained* elsewhere in the transfer list as the response to some other stimulus. The four combinations, 1a, 1b, 2a, and 2b, form a 2 × 2 table where across the top, say, is denoted which kind of response member the transfer-task pair has, new (1) or old (2), and down the side is denoted the transfer-task status of the original (first-task) response, removed (a) or retained (b). This table is shown as Table 2. Conditions 1a and 2b are the usual A-B, A-D and A-B, A-Br paradigms, respectively.

As can be seen from the signed transfer effects listed in Table 2, the A-B, A-Br paradigm involves not only a positive R effect and negative F and B effects, as would be expected from a simple consideration of the three com-

TABLE 2
BREAKDOWN OF TRANSFER EFFECTS
IN THE A-B, A-D AND A-B,
A-Br PARADIGMS

Status of original response	Type of second-list response	
	New (1)	Old (2)
Removed (a)	(A-B, A-D)	
	R: 0	R: +
	F: -	F: -
	B: 0	B: -
Retained (b)		(A-B, A-Br)
	R: 0	R: +
	F: -, -	F: -, -
	B: 0	B: -

ponent effects, but an additional negative F effect. This is because the presence of old, first-task responses tends to elicit backwardly the interfering forward associations acquired in the first task, thus causing these associations to be more resistant to extinction.⁷ An experiment utilizing essentially the design of Table 2 has been conducted by Besch, Thompson, and Wetzel (1962) with results consonant with expectations based on the table.

The third peculiarity attaching to the A-B, A-Br paradigm which has theoretical significance is that transfer in an A-B, A-Br situation must necessarily be a list phenomenon: The re-pairing operation cannot be carried out on a single pair. Transfer theory as discussed in this paper prior to the treatment of the A-B, A-Br is in no way contingent upon a distinction between transfer from a first to a second list and transfer from a first to a second pair.

From the preceding observations, two queries come to mind in quick suc-

⁷ That first-list associations are extinguished during second-list learning has been adequately established. For evidence and discussion on this point, see Briggs (1954), Barnes and Underwood (1959), and Postman (1961, 1962a).

cession. First, since it is clear that the noted peculiarities of the A-B, A-Br paradigm arise from the operation of re-pairing, is it not plausible to suppose that such an operation can define a whole new set of transfer paradigms? And second, since a system of component surfaces clearly serves well to account for transfer phenomena in unrepaired (ordinary) paradigms, is it not reasonable to wonder if an equally serviceable conceptualization might not be devised for the re-paired paradigms?

The set of re-paired paradigms is not difficult to characterize. In the first place, they will be as numerous as the ordinary paradigms because for any ordinary paradigm a re-paired paradigm can be obtained by simply carrying out the re-pairing operation. Repaired paradigms involving completely new transfer-task stimuli and/or responses would not, of course, be distinguishable from their corresponding ordinary paradigms. This means that for such indistinguishable paradigms any system of transfer surfaces designed to handle the re-paired paradigms must coincide with the system of surfaces shown in Figure 1. The loci of coincidence would be along the edges from A-B, C-B to A-B, C-D and from A-B, C-D to A-B, A-D.

Another characteristic of such a system of re-paired paradigms is that the re-paired F and B surfaces must be negative for all positions where first-task forward and backward associations are elicited in the transfer task. The R surface, however, would be the same in both the re-paired and ordinary cases. As far as the present writer knows, the only data pertaining to re-paired transfer situations is that involving the A-B, A-Br paradigm: There have been no studies in which nonzero values of X_s and X_R have been utilized.

As a final remark on the A-B, A-Br paradigm, it is of interest to note that

compared to the A-B, A-D the A-B, A-Br involves (besides its extra negative F effect) the net R-B effect, the opposing positive and negative effects of transferred response availability, and backward associations. This opens up the possibility of expecting *less* negative transfer using the A-B, A-Br paradigm than the A-B, A-D for experimental specifications generating a sufficiently positive net R-B effect. Such a situation would arise whenever either very-low-M response materials were used and/or very few trials of L_1 were given, thus making response learning the predominant first-task process. In fact, using trigrams as responses with L_1 taken to a 3/6 criterion (approximately 3.7 trials), Jung (1962) reports transfer in the A-B, A-Br paradigm to be less negative than that in the A-B, A-D; that is, he reports a positive R-B net effect. At a higher degree of first-task learning (a 6/6 + 5 criterion), the crossover has occurred, and the A-B, A-Br paradigm is more negative than the A-B, A-D. The corresponding role of first-task response M was demonstrated by Merikle and Battig (1963): When first-task responses were low M (CCCs), approximately zero transfer was observed in the A-B, A-D paradigm but considerable positive transfer in the A-B, A-Br; however, as M increased, transfer in the A-B, A-Br paradigm became even more negative than the increasingly negative A-B, A-D. In general, studies using adjectives as responses report more negative transfer in the A-B, A-Br than in the A-B, A-D paradigm (Besch & Reynolds, 1958; Besch, Thompson, & Wetzel, 1962; Kausler & Kanoti, 1963; Keppel & Underwood, 1962; Porter & Duncan, 1953; Postman, 1962b; Twedt & Underwood, 1959), whereas those using trigrams report less (Jung, 1962; Mandler & Heinemann, 1956). Thus

the net R-B effect plays the same role in the A-B, A-Br paradigm as in the A-B, C-B paradigm.

REFERENCES

- ATKINSON, R. C., & ESTES, W. K. Stimulus sampling theory. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology*. Vol. 2. New York: Wiley, 1963. Pp. 121-268.
- BARNES, J. M., & UNDERWOOD, B. J. "Fate" of first-list associations in transfer theory. *Journal of Experimental Psychology*, 1959, 58, 97-105.
- BESCH, N. F., & REYNOLDS, W. F. Associative interference in verbal paired-associate learning. *Journal of Experimental Psychology*, 1958, 55, 554-558.
- BESCH, N. F., THOMPSON, V. E., & WETZEL, A. B. Studies in associative interference. *Journal of Experimental Psychology*, 1962, 63, 342-352.
- BRIGGS, G. E. Acquisition, extinction, and recovery functions in retroactive inhibition. *Journal of Experimental Psychology*, 1954, 47, 285-293.
- BRUCE, R. W. Conditions of transfer of training. *Journal of Experimental Psychology*, 1933, 16, 343-361.
- BUGELSKI, B. R., & CADWALLADER, T. C. A reappraisal of the transfer and retroaction surface. *Journal of Experimental Psychology*, 1956, 52, 360-366.
- BUNCH, M. E., & MCCRAVEN, V. G. The temporal course of transfer in the learning of memory material. *Journal of Comparative Psychology*, 1938, 25, 481-496.
- CARTERETTE, E. C. A replication of free recall and ordering of trigrams. *Journal of Experimental Psychology*, 1963, 66, 311-313.
- CROTHERS, E. J. Paired-associate learning with compound responses. *Journal of Verbal Learning and Verbal Behavior*, 1962, 1, 66-70.
- DALLET, K. M. The transfer surface re-examined. *Journal of Verbal Learning and Verbal Behavior*, 1962, 1, 91-94.
- DEAN, M. G., & KAUSLER, D. H. Degree of first-list learning and stimulus meaningfulness as related to transfer in the A-B, C-B paradigm. *Journal of Verbal Learning and Verbal Behavior*, 1964, 3, 330-334.
- EBBINGHAUS, H. *Memory: A contribution to experimental psychology*. Leipzig: Duncker & Humboldt, 1885. (Trans. by H. A. Ruger) New York: Teacher's College, Columbia Univer., Bureau of Publications, 1913.
- FELDMAN, S. M., & UNDERWOOD, B. J. Stimulus recall following paired-associate learning. *Journal of Experimental Psychology*, 1957, 53, 11-15.
- GAGNÉ, R. M., BAKER, K. E., & FOSTER, H. On the relation between similarity and transfer of training in the learning of discriminative motor task. *Psychological Review*, 1950, 57, 67-79.
- HAMILTON, C. E. The relationship between length of interval separating two learning tasks and performance on the second task. *Journal of Experimental Psychology*, 1950, 40, 613-621.
- HARCUM, E. R. Verbal transfer of overlearned forward and backward associations. *American Journal of Psychology*, 1953, 66, 622-625.
- HOROWITZ, L. M. Free recall and ordering in trigrams. *Journal of Experimental Psychology*, 1961, 62, 51-57.
- JUNG, J. Transfer of training as a function of degree of first-list learning. *Journal of Verbal Learning and Verbal Behavior*, 1962, 1, 197-199.
- JUNG, J. Effects of response meaningfulness (*m*) on transfer of training under two different paradigms. *Journal of Experimental Psychology*, 1963, 65, 377-384.
- KAUSLER, D. A., & KANOTI, G. A. R-S learning and negative transfer effects with a mixed list. *Journal of Experimental Psychology*, 1963, 65, 201-205.
- KEPPEL, G., & UNDERWOOD, B. J. Retroactive inhibition of R-S associations. *Journal of Experimental Psychology*, 1962, 64, 400-404.
- MANDLER, G. From association to structure. *Psychological Review*, 1962, 69, 415-427.
- MANDLER, G., & HEINEMANN, S. H. Effect of overlearning of a verbal response on transfer of training. *Journal of Experimental Psychology*, 1956, 51, 39-46.
- MCGUIRE, W. J. A multiprocess model for paired-associate learning. *Journal of Experimental Psychology*, 1961, 62, 335-347.
- MELTON, A. W. Implications of short-term memory for a general theory of memory. *Journal of Verbal Learning and Verbal Behavior*, 1963, 2, 1-21.
- MERIKLE, P. M., & BATIG, W. F. Transfer of training as a function of experimental paradigm and meaningfulness. *Journal of Verbal Learning and Verbal Behavior*, 1963, 2, 485-488.
- MÜLLER, G. E., & PILZECKER, A. Experimentelle Beiträge zur Lehre vom Gedächtnis.

- niss. *Zeitschrift für Psychologie*, 1900, 1, 1-300.
- MÜLLER, G. E., & SCHUMANN, F. Experimentelle Beiträge zur Untersuchung des Gedächtnisses. *Zeitschrift für Psychologie*, 1894, 6, 81-190, 257-339.
- MURDOCK, B. B., JR. "Backward" associations in transfer and learning. *Journal of Experimental Psychology*, 1958, 55, 111-114.
- NEWTON, J. M., & WICKENS, D. D. Retroactive inhibition as a function of the temporal position of the interpolated learning. *Journal of Experimental Psychology*, 1956, 51, 149-154.
- NOBLE, C. E. Meaningfulness and familiarity. In C. N. Cofer & B. S. Musgrave (Eds.), *Verbal behavior and learning*. New York: McGraw-Hill, 1963. Pp. 76-119.
- OSGOOD, C. E. The similarity paradox in human learning: A resolution. *Psychological Review*, 1949, 56, 132-143.
- PETERSON, L. R., & PETERSON, M. J. Short-term retention of individual verbal items. *Journal of Experimental Psychology*, 1959, 58, 193-198.
- PORTER, L. W., & DUNCAN, C. P. Negative transfer in verbal learning. *Journal of Experimental Psychology*, 1953, 46, 61-64.
- POSTMAN, L. The present status of interference theory. In C. N. Cofer (Ed.), *Verbal learning and verbal behavior*. New York: McGraw-Hill, 1961. Pp. 152-179.
- POSTMAN, L. Retention of first-list associations as a function of the conditions of transfer. *Journal of Experimental Psychology*, 1962, 64, 380-387. (a)
- POSTMAN, L. Transfer of training as a function of experimental paradigm and degree of first-list learning. *Journal of Verbal Learning and Verbal Behavior*, 1962, 1, 109-118. (b)
- SCHULZ, R. W., & MARTIN, E. Aural paired-associate learning: Stimulus familiarization, response familiarization, and pronunciability. *Journal of Verbal Learning and Verbal Behavior*, 1964, 3, 139-145.
- SCHULZ, R. W., & TUCKER, I. F. Stimulus familiarization and length of the anticipation interval in paired-associate learning. *Psychological Reports*, 1962, 12, 341-344.
- SHEPARD, R. N. Comments on Professor Underwood's paper. In C. N. Cofer & B. S. Musgrave (Eds.), *Verbal behavior and learning*. New York: McGraw-Hill, 1963. Pp. 48-70.
- SPIKER, C. C. Associative transfer in verbal paired-associate learning. *Child Development*, 1960, 31, 73-87.
- THORNDIKE, E. L. *The fundamentals of learning*. New York: Columbia University Press, 1932.
- TWEDT, H. M., & UNDERWOOD, B. J. Mixed vs. unmixed lists in transfer studies. *Journal of Experimental Psychology*, 1959, 58, 111-116.
- UNDERWOOD, B. J. *Experimental psychology*. New York: Appleton-Century-Crofts, 1949.
- UNDERWOOD, B. J., RUNQUIST, W. N., & SCHULZ, R. W. Response learning in paired-associate lists as a function of intra-list similarity. *Journal of Experimental Psychology*, 1959, 58, 70-78.
- UNDERWOOD, B. J., & SCHULZ, R. W. *Meaningfulness and verbal learning*. New York: Lippincott, 1960.
- WIMER, R. Osgood's transfer surface: Extension and test. *Journal of Verbal Learning and Verbal Behavior*, 1964, 3, 274-279.
- YOUNG, R. K., & UNDERWOOD, B. J. Transfer in verbal materials with dissimilar stimuli and response similarity varied. *Journal of Experimental Psychology*, 1954, 47, 153-159.

(Received July 13, 1964)

SOME STIMULUS DIMENSIONS OF ROTATING SPIRALS

THOMAS R. SCOTT

AND

J. H. NOLAND

VA Hospital, Columbia, South Carolina

University of South Carolina

The perceptible motion of rotating spiral lines can be analyzed into 3 components: motion normal to the line, rotational motion, and radial motion. General equations for these 3 components have been derived. Specific formulas for finding the 3 components for Archimedes, logarithmic, and hyperbolic spirals have been given in terms of distance from the center, speed of rotation, and the constants associated with these spirals. In addition, 3 "special" spirals have been derived which have, respectively, the properties that the normal, radial, and rotational motions are constant for all distances from the center of rotation greater than a minimum distance. Possible applications have been suggested.

Several new methods for measuring aftereffect of seen motion have recently been developed (Carlson, 1962; Scott, Bragg, & Smarr, 1963; Scott, Jordan, & Powell, 1963; Scott & Medlin, 1962; Scott & Powell, 1963; Singer, 1959; Taylor, 1963), and recent physiological findings (Hubel, 1963; Maturana, 1963) have given promise of an imminent increase in the understanding of these effects. Eliciting motions of three principal types have been used: linear, rotary, and spiral. The first two types of motion are fairly simple, and little difficulty has been encountered in the quantification of the stimulus dimensions. However, in the case of the spiral, whose motion is somewhat more complex, no detailed treatment of the stimulus dimensions appears to have been worked out. Any precise treatment of aftereffect produced by rotating spirals depends crucially upon the precision of knowledge about the stimulus dimensions. The present discussion attempts to provide an analysis of those stimulus dimensions of rotating spirals which are of greatest potential interest to the investigator of visual motion aftereffect.

The first puzzle which must be clarified is the apparent mystery asso-

ciated with the so-called "spiral effect" or "spiral illusion" (see Stern, 1959). Some students have found it hard to understand how a rotating spiral can appear to expand or contract when it is "really" just turning. It is obvious that a tightly wound spiral with many turns will appear to have less radial motion than a more open one with fewer turns rotated at the same speed. Consider, however, a spiral with zero turns. Such a "spiral" is really just a "spoke." Rotation of such a stimulus will elicit only rotational aftereffect. Consider now a spiral with, say, a single turn. This figure will look like a "curved spoke." It can be easily observed that this figure will elicit two kinds of aftereffect at the same time: rotational and radial. Somehow, then, as the number of turns of the spiral becomes less, the perceived motion and its aftereffect undergo a gradual change from almost purely radial to entirely rotational. The stimulus object, of course, has only a rotational motion. How, then, can the transformation from a rotary motion into a composite of rotary and radial motions take place?

To answer this question, it is helpful to consider the conditions necessary

for the perception of motion per se. No homogeneous contour on a homogeneous field can be seen to move if the successive positions of the motion do not change the visual pattern received by the eye. For example, a circle rotated about its center, or a straight line moved along its axis, will be self-congruent (see Platt, 1958) for all positions for the types of motion involved, and no motion will be perceived in such cases. These considerations suggest the axiom that *the only perceptible motion of a contour is that component of its motion which is normal to the contour*. If the contour is curved, this component will be normal to the curve at the point whose perceptible motion is to be evaluated. When seen in this light, the apparent expansion or contraction of a rotating spiral is not an illusion at all but is the direct consequence of the perceptually effective motion of the spiral lines. The "latency of the spiral effect" as studied by Stern (1959) is, then, no more than the latency for motion perception under the particular conditions used. Further analysis can now be undertaken.

The following questions need to be answered: What is the exact speed of the perceptible motion of a given rotating spiral at a given distance from the center of rotation? ("Perceptible motion" here is used in the sense of the axiom stated above.) What is the direction of this motion? How much of the perceptible motion can be represented as radial and how much rotational? What types of spirals will produce what functional relationships between radial motion, rotational motion, and distance from the center?

In order to answer these questions it is necessary to state in mathematical form what is meant by a *spiral*. A geometric analysis can then be under-

taken in order to discover the consequences of the perceptible-motion axiom as it applies to particular types of spirals.

For the purposes of this discussion, a spiral is defined as any monotonic function of the form

$$\rho = f(\theta),$$

where ρ and θ are polar coordinates. There are thus as many different kinds of spirals as there are monotonic functions of θ . Consider the spiral OPX (see Figure 1) with P any point on the spiral. APB is tangent to the spiral at the point P. The objective rotational velocity is represented by the vector \overline{PZ} , perpendicular to the radius vector ρ . If rotational velocity is represented by ω revolutions per unit of time,

$$\overline{PZ} = 2\pi\omega\rho.$$

Let PN be perpendicular to APB. The vector \overline{PQ} is the projection of \overline{PZ} on PN. It is proposed that the vector \overline{PQ} represents the velocity of the perceptible component of the motion of the spiral line at the point P. Let $\angle APO = \psi$. Since $\angle NPO = 90^\circ + \psi$ and $\angle ZPO = 90^\circ$, $\angle ZPQ = \psi$, and the length of vector \overline{PQ} is therefore given by

$$\overline{PQ} = 2\pi\omega\rho \cos \psi. \quad [1]$$

Equation 1 represents the total perceptible motion which will be called normal motion since its direction is normal to the curve. It can, of course, be resolved into a radial component and a rotational component. These are shown by the vectors \overline{PE} and \overline{PK} , respectively. The length of \overline{PE} , or the speed of radial motion, is given by

$$\begin{aligned} \overline{PE} &= \overline{PQ} \sin \psi, \\ &= 2\pi\omega\rho \cos \psi \sin \psi. \end{aligned} \quad [2]$$

Equation 1 becomes

$$\begin{aligned}\overline{PQ} &= 2\pi\omega\rho \cos \arctan \rho/a, \\ &= \frac{2a\pi\omega\rho}{\sqrt{\rho^2 + a^2}}.\end{aligned}$$

Equation 2 becomes

$$\begin{aligned}\overline{PE} &= 2\pi\omega\rho \cos \arctan \rho/a \\ &\quad \sin \arctan \rho/a, \\ &= \frac{2a\pi\omega\rho^2}{\rho^2 + a^2}.\end{aligned}$$

Equation 3 becomes

$$\begin{aligned}\overline{PK} &= 2\pi\omega\rho \cos^2 \arctan \rho/a, \\ &= \frac{2a^2\pi\omega\rho}{\rho^2 + a^2}.\end{aligned}$$

Similar substitutions of Equations 6 and 7 yield corresponding expressions for the logarithmic and hyperbolic spirals. These formulas are summarized in Table 1.

TABLE 1

SUMMARY OF FORMULAS FOR THE INDICATED COMPONENTS OF PERCEIVED MOTION FOR THREE KINDS OF SPIRALS AS A FUNCTION OF DISTANCE FROM THE CENTER (ρ) AND SPEED OF ROTATION (ω)

Spiral formula	Normal motion (\overline{PQ})	Radial motion (\overline{PE})	Rotational motion (\overline{PK})
Archimedes: $\rho = a\theta$	$\frac{2a\pi\omega\rho}{\sqrt{\rho^2 + a^2}}$	$\frac{2a\pi\omega\rho^2}{\rho^2 + a^2}$	$\frac{2a^2\pi\omega\rho}{\rho^2 + a^2}$
Logarithmic: $\rho = e^{b\theta}$	$\frac{2b\pi\omega\rho}{\sqrt{b^2 + 1}}$	$\frac{2b\pi\omega\rho}{b^2 + 1}$	$\frac{2b^2\pi\omega\rho}{b^2 + 1}$
Hyperbolic: $\rho\theta = c$	$\frac{2\pi\omega\rho^2}{\sqrt{\rho^2 + c^2}}$	$\frac{2c\pi\omega\rho^2}{\rho^2 + c^2}$	$\frac{2\pi\omega\rho^3}{\rho^2 + c^2}$

Figure 2 represents graphically the normal, radial, and rotational motions as functions of distance from the center of an Archimedes spiral. It can be seen by inspection of the

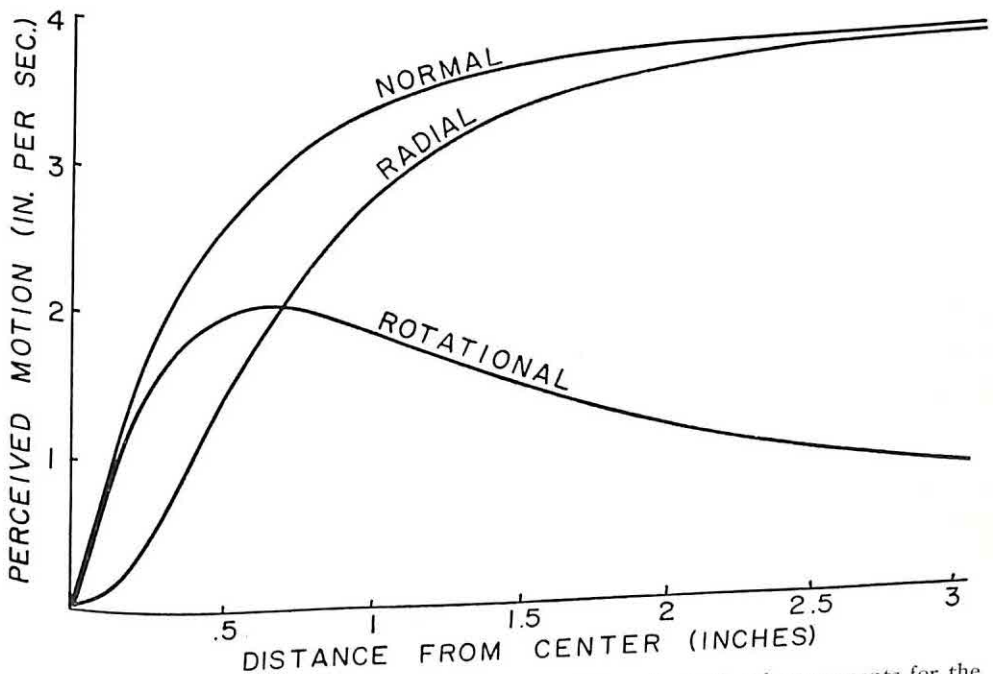


FIG. 2. Graphic representation of the normal, radial, and rotational components for the Archimedes spiral as a function of distance from the center. ($a = .64$, $\omega = 1$ revolution per second.)

formula and the curves in Figure 2 that the normal motion increases with negative acceleration and approaches $2a\pi\omega$ asymptotically. The radial motion, although somewhat less than the normal motion, approaches the same asymptote. Rotational motion reaches a maximum and thereafter approaches zero. To find the maximum, the original formula,

$$\overline{PK} = \frac{2a^2\pi\omega\rho}{\rho^2 + a^2},$$

can be differentiated with respect to ρ :

$$\frac{d\overline{PK}}{d\rho} = -2a^2\pi\omega\rho[a^2 + \rho^2]^{-2}[2\rho] + 2a^2\pi\omega[a^2 + \rho^2]^{-1}.$$

Simplifying and setting equal to zero:

$$\frac{2a^2\pi\omega[a^2 + \rho^2] - 4a^2\pi\omega\rho^2}{[a^2 + \rho^2]^2} = 0.$$

It can be seen that this function approaches zero as ρ (and therefore the denominator) approaches infinity. The desired maximum, however, is where the numerator equals zero as follows:

$$2a^2\pi\omega[a^2 + \rho^2 - 2\rho^2] = 0.$$

But

$$2a^2\pi\omega \neq 0.$$

Therefore

$$a^2 - \rho^2 = 0 \\ \rho = a.$$

Interestingly, then, the maximum rotational motion for the Archimedes spiral is obtained exactly a units from the center of rotation. Substituting this value of ρ into the original formula:

$$\overline{PK}_{\max} = \frac{2a^2\pi\omega a}{a^2 + a^2} = \pi\omega a.$$

The maximum rotational motion is thus exactly half the limiting normal-

and radial-motion rates. The Archimedes spiral, then, in addition to being easy to draw, would be the stimulus of choice in experiments where relatively uniform radial motion is desirable, and where little rotational motion is wanted—provided that the experimenter studies motion substantially more than a units from the center of rotation.

The logarithmic spiral has different but equally interesting properties. Inspection of Figure 3 and the second line of Table 1 shows that all three types of motion increase without limit as linear functions of distance from the center of rotation. The logarithmic spiral has been a favorite geometric figure among mathematicians for many years because of its numerous interesting properties (see Turnbull, 1956; Weyl, 1956), and it is not surprising that it also has interesting perceptual properties when rotated. A rotating logarithmic spiral filling the visual field would be visually similar to an approaching or receding object with parallax and accommodation artificially removed from the situation. As a laboratory stimulus it would also have the advantage that "approach" or "recession" could be continued for indefinite lengths of time without either running into the stimulus or having the stimulus-observer distance exceed reasonable values. It should be possible to demonstrate a very curious perceptual ambiguity with this kind of spiral. Under appropriately reduced cues, an observer should be unable to tell whether the logarithmic spiral is changing its distance from him or whether it is simply rotating! From the expressions in Table 1, it can easily be seen that the ratio of radial to rotational motion of a logarithmic spiral is everywhere $1/b$. The value of b used in constructing Figure 3 was

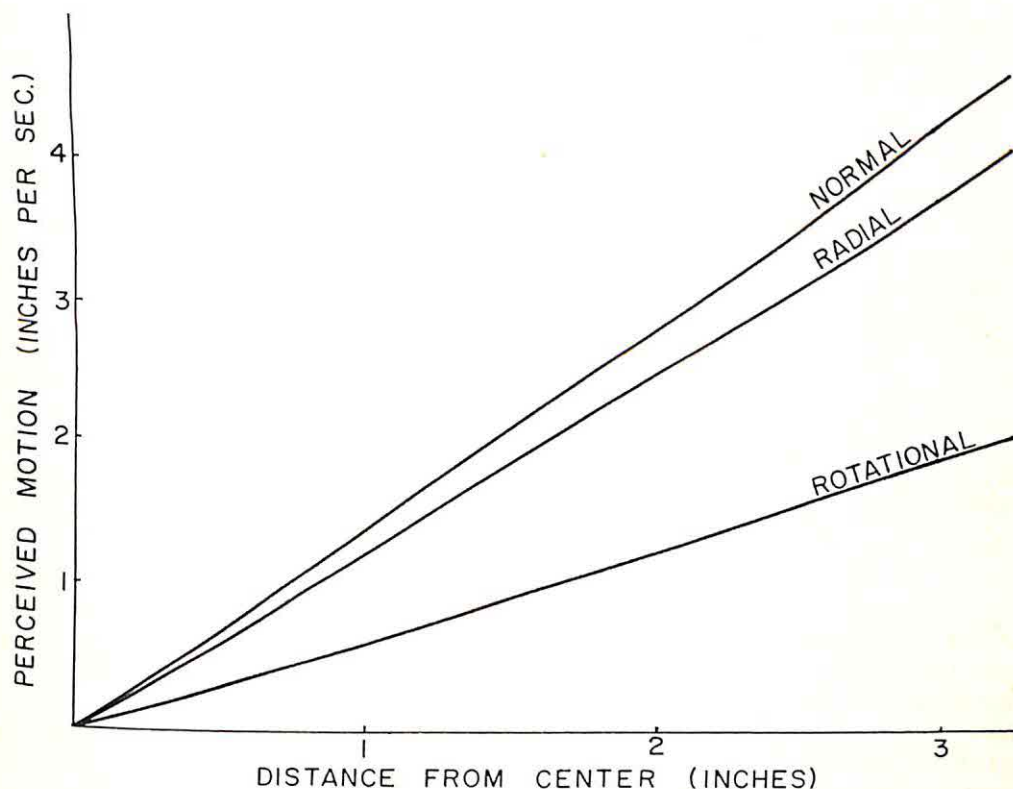


FIG. 3. Graphic representation of the normal, radial, and rotational components for a logarithmic spiral. ($b = .5$, $\omega = 1$ revolution per second.)

chosen to produce twice as much radial as rotational motion. The ratio of radial to rotational motion can, of course, be given any desired value by proper selection of the value of b .

With both the Archimedes and the logarithmic spirals, it is possible to increase the number of turns by increasing the size of the spiral. This is impossible with the hyperbolic spiral, however, because the spiral approaches a straight line c units from the origin and parallel to the reference vector. For the hyperbolic spiral, both the normal motion and the rotational motion increase without limit as positively accelerated functions of distance from the center of rotation (see Figure 4). The radial motion becomes asymptotic to $2c\pi\omega$ thus seeming to mimic the Archimedes.

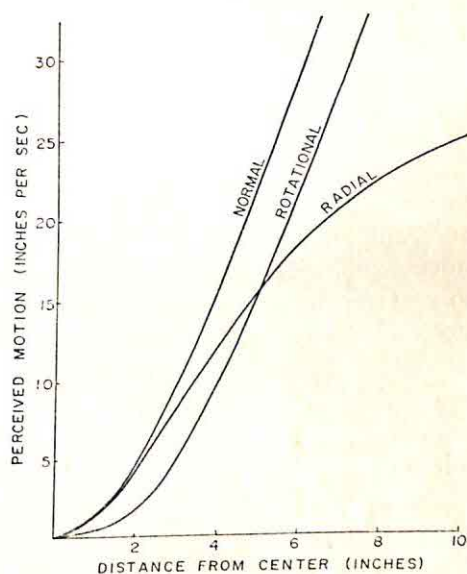


FIG. 4. Graphic representation of the normal, rotational, and radial components for a hyperbolic spiral. ($c = 5$, $\omega = 1$ revolution per second.)

However, at or beyond a value of ρ for which the radial component is close to its asymptote, the rotational component has become so large that radial motion is only a negligible part of the total normal motion. The hyperbolic spiral would thus seem to be a poor choice for most studies of visually perceived motion.

All previous studies of spiral aftereffect have used one of the three types of spirals discussed above. For specific experimental purposes, however, it may prove advantageous to construct special spirals which have particularly desirable attributes for studying specific problems. In particular, it may be of interest to design spirals for which certain components of motion are constant for all distances (greater than a minimum) from the center. Three such spirals will now be developed. The method will be to substitute expressions for $\sin \psi$ and $\cos \psi$ in terms of ρ and ρ' into Equations 1, 2, and 3. The resulting equations will then be treated as differential equations with the vectors \overline{PQ} , \overline{PE} , and \overline{PK} treated as constants.

The first special spiral might be called the "isonormal" spiral. This is a spiral which, when rotated, has the same normal motion at all distances from the center greater than a minimum distance. This minimum distance would be where the desired normal motion equals the objective motion. In symbolic terms, it would be where $\overline{PZ} = \overline{PQ}$. For distances less than this, the isonormal spiral is undefined. From Equation 4a it follows that

$$\cos \psi = \frac{\rho'}{\sqrt{\rho^2 + \rho'^2}}, \quad [8]$$

and

$$\sin \psi = \frac{\rho}{\sqrt{\rho^2 + \rho'^2}}. \quad [9]$$

To derive a differential equation for the isonormal spiral, Equation 1 can be solved for $\cos \psi$:

$$\cos \psi = \frac{\overline{PQ}}{2\pi\omega\rho}.$$

Substituting from Equation 8,

$$\frac{\rho'}{\sqrt{\rho^2 + \rho'^2}} = \frac{\overline{PQ}}{2\pi\omega\rho}.$$

Squaring both sides, rearranging into standard form, and writing $\frac{d\rho}{d\theta}$ for ρ' results in

$$4\pi^2\omega^2\rho^2 \left[\frac{d\rho}{d\theta} \right]^2 - \overline{PQ}^2\rho^2 - \overline{PQ}^2 \left[\frac{d\rho}{d\theta} \right]^2 = 0, \quad [10]$$

which is the desired differential equation. The solution can be expressed as follows:¹

$$\theta = \sqrt{\left[\frac{2\pi\omega\rho}{\overline{PQ}} \right]^2 - 1} - \arctan \sqrt{\left[\frac{2\pi\omega\rho}{\overline{PQ}} \right]^2 - 1}.$$

Figure 5 shows the first two turns of the isonormal spiral. For larger values of ρ the spiral is approximately an Archimedes spiral whose formula is

$$\rho = \left[\frac{\overline{PQ}}{2\pi\omega} \right] \left[\theta + \frac{\pi}{2} \right].$$

The second special spiral has the property that the radial motion is constant for all distances from the

¹ Specifications for isonormal, isorotational, and isoradial spirals have been deposited with the American Documentation Institute. Order Document No. 8333 from ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington, D. C. 20540. Remit in advance \$1.75 for microfilm or \$2.50 for photocopies and make checks payable to: Chief, Photoduplication Service, Library of Congress.

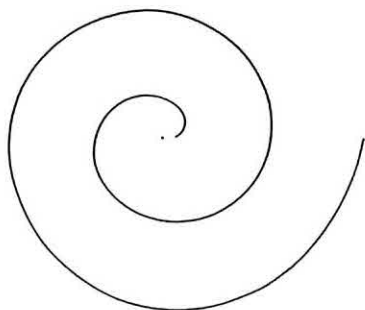


FIG. 5. First two turns of the isonormal spiral. (Rotation of this figure produces the same total perceptible motion—motion normal to the curve—at all distances from the center. The center is indicated by the dot.)

center greater than a minimum. The minimum for the isoradial spiral is the point where the objective motion is twice the desired radial component. In symbols, $\overline{PZ} = 2 \overline{PE}$. As proof of the latter, it can be readily seen (see Figure 1) that the maximum radial component for any given value of \overline{PZ} and ρ will be where $\psi = 45^\circ$. Triangle ZPQ will be an isosceles right triangle, K will bisect \overline{PZ} , and \overline{PK} will equal \overline{QK} . Since $\overline{QK} = \overline{PE}$, $\overline{PZ} = 2 \overline{PE}$.

To find the differential equation, substitute from 8 and 9 into 2 as follows:

$$\overline{PE} = 2\pi\omega\rho \frac{\rho'}{\sqrt{\rho^2 + \rho'^2}} \cdot \frac{\rho}{\sqrt{\rho^2 + \rho'^2}}$$

which, in standard form, becomes:

$$2\pi\omega\rho^2 \left[\frac{d\rho}{d\theta} \right] - \overline{PE}\rho^2 - \overline{PE} \left[\frac{d\rho}{d\theta} \right]^2 = 0. \quad [11]$$

The solution is:

$$\theta + 1 = \frac{\pi\omega\rho}{\overline{PE}} \pm \left[\sqrt{\left[\frac{\pi\omega\rho}{\overline{PE}} \right]^2 - 1} - \arctan \sqrt{\left[\frac{\pi\omega\rho}{\overline{PE}} \right]^2 - 1} \right].$$

Figure 6 shows the first two turns of the isoradial spiral (using positive values of the bracketed function). For larger values of ρ this equation, too, is approximately that of an Archimedes spiral. This time the approximation formula is:

$$\rho = \frac{\overline{PE}}{2\pi\omega} \left[\theta + 1 + \frac{\pi}{2} \right].$$

The third and last special spiral is the isorotational spiral which has the same speed of rotational motion at all distances from the center, greater than a minimum. As in the isonormal spiral, the minimum is the point where the desired component equals the objective motion, or in this case, where $\overline{PZ} = \overline{PK}$. For values of ρ less than this, the isorotational spiral is undefined. Solving Equation 3 for $\cos^2 \psi$,

$$\cos^2 \psi = \frac{\overline{PK}}{2\pi\omega\rho}.$$

Substituting from Equation 8,

$$\frac{\rho'^2}{\rho^2 + \rho'^2} = \frac{\overline{PK}}{2\pi\omega\rho},$$

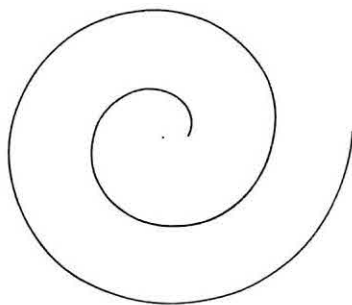


FIG. 6. First two turns of the isoradial spiral. (Rotation of this figure produces the same perceptible speed of radial motion at all points on the curve. This figure is to be contrasted with Figure 5. The final value of ρ is the same for both figures.)

and rearranging into standard form yields

$$2\pi\omega\rho\left[\frac{d\rho}{d\theta}\right]^2 - \overline{PK}\rho^2 - \overline{PK}\left[\frac{d\rho}{d\theta}\right]^2 = 0, \quad [12]$$

which is the differential equation for the isorotational spiral. The solution is quite similar to that for Equation 10:

$$\theta = 2\sqrt{\frac{2\pi\omega\rho}{\overline{PK}}} - 1 - 2\arctan\sqrt{\frac{2\pi\omega\rho}{\overline{PK}}} - 1.$$

Figure 7 shows the first two turns of the isorotational spiral. For larger values of ρ , the spiral has approximately the form

$$\rho = \frac{\overline{PQ}}{8\pi\omega} [\theta + \pi]^2.$$

Analysis of the stimulus dimensions of rotating spirals, then, turns out to be somewhat more involved than one might at first expect. The student of perception might be tempted to bypass these complexities by employing simpler stimuli. Several investigators have chosen simpler stimuli on the grounds that the spiral introduced

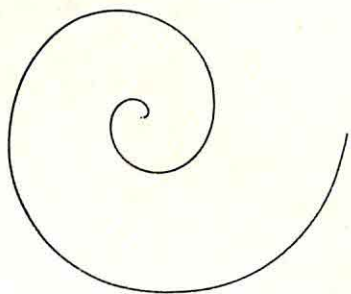


FIG. 7. First two turns of the isorotational spiral. (When this figure is rotated, the perceptible rotational component is the same at all distances from the center. The final value of ρ is equal to that of Figures 5 and 6.)

unnecessary complexities into the experimental situation. Prior to the present analysis, such arguments were, of course, quite cogent due to the unknown character of these complexities. Now that an analysis has been made, however, there are some rather compelling advantages to the use of rotating spirals. Motion aftereffect was first studied with linear motion, as in the familiar waterfall illusion. Most theories of aftereffect at that early date had to do with eye movements and eye-muscle innervations. Plateau (1850) disproved the universality of such theories precisely by introducing the rotating spiral as a stimulus. The rotating spiral with a central fixation point is a superior stimulus in any experiment where eye movements might be an important control variable. The method of measuring rate-of-motion aftereffect introduced by Cords and Brucke (1907) and extended by Singer (1959); Scott, Bragg, and Smarr (1963); Scott, Jordan, and Powell (1963); Taylor (1963); and perhaps others, in which a test stimulus moves slowly in the same direction as the inspection stimulus so as to cancel the aftereffect, presents peculiar difficulties in waterfall-illusion experiments. In such experiments a pattern of moving stripes typically is seen during the inspection period inside some kind of frame, usually rectangular. The test stimulus, then, is seen inside the same or a similar frame. The real motion of such a test stimulus can be perceived by comparison of point correspondence from the stimulus to the frame and by distance observations made relative to the frame. A possible way of avoiding these undesired cues is the elimination of the frame. Here the inspection stimulus and the test stimulus are seen against a field of total darkness. The difficulty with

this solution is that during the test period the well-known phenomenon of autokinetic movement may introduce unknown and uncontrolled amounts of error which have in addition the disadvantage of being highly susceptible to suggestion and preconception on the part of the subject. The use of a rotating stimulus having radial lines drawn on it, for example in Wohlgenuth's (1911) studies, has many of the advantages associated with the spiral. Indeed, the present analysis shows that such stimuli are merely special cases of the spiral. More complex rotating patterns, as in Taylor's "swirling lines," have the advantages of rotating stimuli in general, but do not lend themselves to the study of more complex problems such as those associated with differential aftereffects toward and away from the fovea.

Different investigators have used spirals differing in type, pitch, number of "throws," proportion of black and white areas, diameter, and speed of rotation. Viewing distances have been different between experiments. Aftereffect while viewing the same pattern stationary has been timed. Aftereffect while viewing some other pattern has been timed. In Scott's experiments, only the radial component of aftereffect was measured, even though rotational aftereffect was certainly present. In Taylor's experiments only rotational aftereffect was measured; even though the possibility that his swirling lines may elicit other directions of aftereffect remains untested. Many of these experiments done by different investigators in different parts of the world are directed at fundamental questions regarding motion perception. When conflicts appear in the data of these diverse studies, some exact method of quantifying the stimulus conditions is

needed. For example, in the studies of Price and Deabler (1955), a 920° Archimedes spiral 6 inches in diameter rotating at 78 and 100 revolutions per minute (rpm) was used. Viewing distance was 8 feet. A study by Scott, Bragg, and Smarr (1963) which differed substantially in result from that of Price and Deabler involved a 1440° Archimedes spiral 8 inches in diameter rotated at 160 rpm and viewed at a distance of 6 feet. In Price and Deabler's study, the spiral itself served as the test stimulus. The aftereffect, then, was presumably aftereffect resulting from the "normal motion" of their spiral. In the study by Scott, Bragg, and Smarr, only the radial aftereffect was measured. This aftereffect presumably resulted from the radial component of the particular spiral involved.

In comparing these two studies, a precise understanding of the difference between the two stimulus conditions is essential before intelligent hypotheses can be made regarding the difference in outcome. Using the method derived here, the normal motion at the edge of the spiral in Price and Deabler's study was 70 minarcs per second for the 100-rpm speed and 54 minarcs per second for the 78-rpm speed. In Scott, Bragg, and Smarr's experiment, the maximum normal motion at the edge of the spiral was 126 minarcs per second. The radial component of this motion was also 126 minarcs per second to the nearest minute of arc. In this latter study, aftereffect was measured at a distance of 1.5 inches from the center of the spiral. Even at this distance, the radial and normal components were still 126 minarcs per second to the nearest minute of arc. These results show that the speed of the eliciting stimulus in the Scott, Bragg, and Smarr experiment was from 80%

to 133% greater than that used by Price and Deabler.

A recent study by Powell (1962) investigated the differential rate of radial aftereffect measured at different distances from the fovea. An Archimedes spiral was used in this study. An initial assumption was made that since the formula for an Archimedes spiral is $\rho = a\theta$ and since the rate of change of ρ with respect to θ is the constant a , that different regions of the retina were being stimulated with the same speed of motion. The present analysis reveals that this was an erroneous assumption. Fortunately in this case, the combination of spiral and range of distances from the fovea chosen by Powell were such that the erroneous assumption was actually a close approximation to the result which would be obtained by the more exact method developed here. This example, however, reveals the necessity of such an analysis for the more detailed study of the dynamics of motion-aftereffect phenomena.

If the theoretical model presented here has validity, it should be possible to show similarities between results obtained by different investigators using similar eliciting speeds of motion. Similar results might be expected in comparing linear-motion aftereffect to spiral aftereffect. Data exist in the literature which make possible comparisons of this kind. From the available data, three studies have been selected for comparison. The first of these was done by Granit in 1928. In Granit's study, the eliciting stimulus was a rotating drum having parallel black-on-white bands drawn on it. The drum was seen through a square window, and the objective speed of the bands was 10 centimeters per second. Granit measured the duration of the aftereffect resulting from this virtually linear

motion at a number of different viewing distances. He was ostensibly studying the effects of retinal area on duration of the aftereffect. Although in his preliminary discussion of the problem he was obviously aware that the speed of motion of the stimulus pattern across the retina also varied with viewing distance, he seemed to ignore this in his subsequent discussion of results. For present purposes, Granit's results will be treated as if speed of the eliciting motion were the only variable, but it should not be forgotten that size of retinal area was also varied in his experiment. Figure 8 shows the average duration of aftereffect obtained by Granit as a function of speed of eliciting motion (solid circles, solid line). The speeds of stimulation are expressed in minarcs per second based on a simple trigonometric transformation of the linear motion of the bands on Granit's drum into units of visual angle. As can be seen, the aftereffect increased from stimulating speeds of about 72 minarcs per second, reaching a maximum at approximately 120 minarcs per second and thereafter decreasing. Granit also varied the brightness of the stimulus and the degree of dark adaptation, but the results from all these different conditions have been pooled as have the results from the three subjects he used in his experiment. It should be noted that the pooling of all these variables might be expected to have some adverse effects on any attempt to compare his results with others. This fact will make any similarity between his results and other results all the more striking.

Scott (1962) carried out a normative study in which eight different Archimedes spirals were used. These spirals varied in pitch and number of throws. The measure of aftereffect used was the rate of expansion or

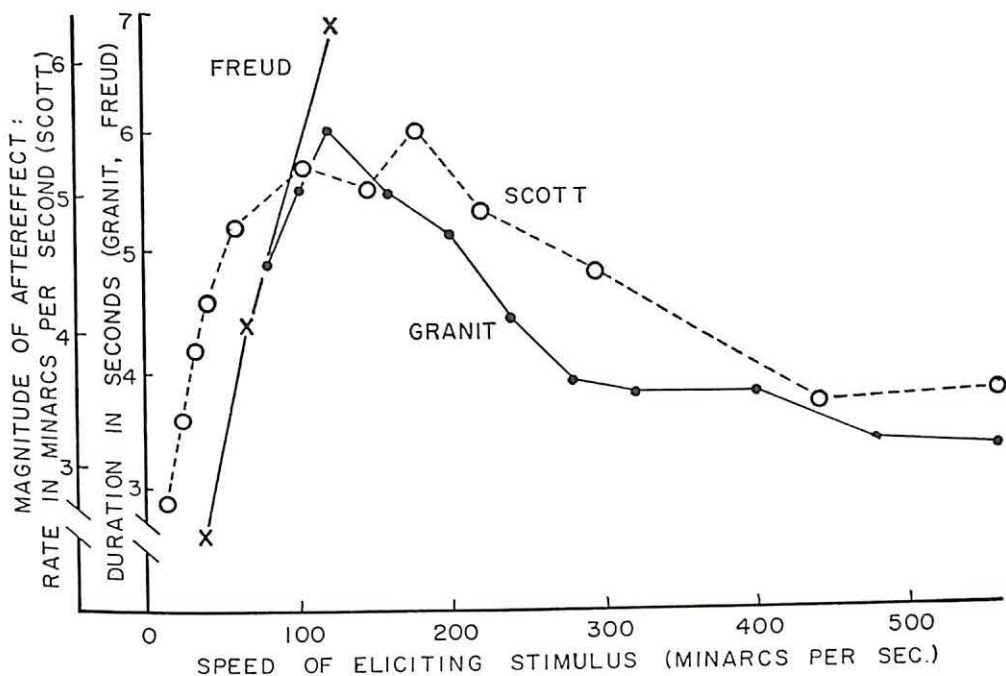


FIG. 8. A comparison of the magnitude of the aftereffect obtained by Granit, Scott, and Freud (ordinate) with the results from all three experiments plotted on a common abscissa: the speed of eliciting motion in minarcs per second using formulas developed here. (Freud's and Granit's results are expressed in seconds of duration of the aftereffect. Scott's data, originally expressed in minarcs per second, are here artificially adjusted so that the highest point is graphed at the highest point of Granit's data.)

contraction of a test stimulus which subjectively cancelled the aftereffect. There would seem to be little hope that a similarity might be found between the results of Scott's study and those of Granit's. The two experiments differed in several respects. First, the eliciting motions were quite different. Second, in Granit's study the speed of stimulation was controlled by changing the viewing distance; in Scott's study the speed was controlled by varying the speed of rotation of the spirals and the pitch of the spirals. Several other variables were involved in both Granit's and Scott's studies. The measure of aftereffect used was different. In spite of these many differences, the curve in Figure 8 which summarizes Scott's results (open circles, dashed lines)

bears a close similarity to the results of Granit. It must be pointed out, however, that since the two measures of aftereffect were not directly comparable, the highest point of Scott's data has been artificially adjusted to match the highest point in Granit's data.

The third study is a recent one by Freud (1962). Freud was also interested in the effect of varying the size of the stimulated retinal area, and, like Granit, he varied the area by manipulating the viewing distance. He used an Archimedes spiral and a constant speed of rotation. He thus committed the same error as did Granit in 1928: ignoring the fact that changing the distance also changes the speed of stimulation. Freud's measure of aftereffect was duration.

Computation of his speeds of eliciting motion and pooling his data across all other conditions results in the data shown in Figure 8 (Xs, solid line). Unfortunately, the range of speeds used by Freud was very narrow, but within that range the results have the same sense as those obtained by Granit and Scott. That is, the aftereffect increases for stimulating speeds between 30 and 132 minarcs per second. Other variables in Freud's study were the location of the stimulated area near or far from the fovea and whether aftereffect was observed monocularly or interocularly. The presence of these other variables fails to obliterate the similarity of results. Thus far, the comparison of various results suggests that motion aftereffect is surprisingly predictable from a knowledge of the speed of the eliciting motion even though other variables are known to influence it.

The availability of the isonormal, isoradial, and isorotational spirals suggests the possibility of some new dynamic studies of radial and rotational aftereffect. Do the responding mechanisms of the eye and brain behave in the Pythagorean manner suggested by the geometric model presented here? Are radial aftereffect and rotational aftereffect independent? Most studies of motion aftereffect have used relatively uniform speeds of stimulation in different areas of the retina (as in the linear-motion studies and the Archimedes-spiral studies). Calculation of stimulating speeds for the *logarithmic* spirals used in two recent experiments (Pickersgill & Jeeves;² Singer, 1959) suggest that under these nonuniform stimulus conditions unusually long durations of aftereffect are encountered. No doubt a number of new experiments will be

devised to test various aspects of the "perceptible-motion axiom" and its consequences.

The equations developed in this article, then, have two immediate applications. The first is to clarify comparisons between the findings of different investigators; the second is in the design of new experiments. Investigators armed with the results of the present analysis should be able not only to resolve seemingly contradictory outcomes from different laboratories but also to choose stimulus materials for specific experimental purposes in a way much more precise than was previously possible.

REFERENCES

- CARLSON, V. R. Adaptation in the perception of visual velocity. *Journal of Experimental Psychology*, 1962, 64, 192-197.
- CORDS, R., & BRUCKE, E. T. Über die Geschwindigkeit des Bewegungsnachbildes. *Pflügers Archiv für die Gesamte Physiologie*, 1907, 119, 54-76.
- FREUD, S. L. *A study of the physiological mechanisms underlying the spiral aftereffect*. (Doctoral dissertation, University of Connecticut) Arlington, Va.: Armed Services Technical Information Agency, 1962, AD No. 274366.
- GRANIT, R. On inhibition of the aftereffect of seen movement. *British Journal of Psychology*, 1928, 19, 147-157.
- GRANVILLE, W. A., SMITH, P. F., & LONGLEY, W. R. *Elements of the differential and integral calculus*. (Rev. ed.) Boston: Ginn, 1941. P. 124.
- HUBEL, D. H. The visual cortex of the brain. *Scientific American*, 1963, 209(5), 54-62.
- MATURANA, H. R., & FRENK, S. Directional movement and horizontal edge detectors in the pigeon retina. *Science*, 1963, 142, 977-979.
- PEIRCE, B. O., & FOSTER, R. M. *A short table of integrals*. (4th ed.) Boston: Ginn, 1956.
- PLATEAU, J. Vierte Notiz über eine neue sonderbare Anwendung des Verweilens der Eindrücke auf der Netzhaut. *Poggendorff's Annalen der Physik Chemie*, 1850, 80, 287-292.
- PLATT, J. R. Functional geometry and the determination of pattern in mosaic recep-

² Mary J. Pickersgill and M. A. Jeeves, personal communication, 1964.

- tors. In H. P. Yockey (Ed.), *Symposium on information theory in biology*. New York: Pergamon Press, 1958. Pp. 371-398.
- POWELL, D. A. Spiral aftereffect rate as a function of the direction of rotation of the spiral and size of the test stimulus. Unpublished master's thesis, University of South Carolina, 1962.
- PRICE, A. C., & DEABLER, H. L. Diagnosis of organicity by means of the spiral aftereffect. *Journal of Consulting Psychology*, 1955, 19, 299-302.
- SCOTT, T. R. Summary of principal findings on a visual aftereffect of motion and brain pathology. *Newsletter for Research in Psychology*, August 1962, 38-46.
- SCOTT, T. R., BRAGG, R. A., & SMARR, R. G. Brain damage diagnosis with the MMG. *Journal of Consulting Psychology*, 1963, 27, 45-53.
- SCOTT, T. R., JORDAN, A. E., & POWELL, D. A. Does visual aftereffect of motion add algebraically to objective motion of the test stimulus? *Journal of Experimental Psychology*, 1963, 66, 500-505.
- SCOTT, T. R., & MEDLIN, R. E. Psychophysical measurement of the spiral aftereffect. *American Journal of Psychology*, 1962, 75, 319-321.
- SCOTT, T. R., & POWELL, D. A. Measurement of a visual motion aftereffect in the rhesus monkey. *Science*, 1963, 140, 57-59.
- SINGER, J. A. The relationship between extent of aftereffect and speed of a rotated spiral. Unpublished doctoral dissertation, Boston University, 1959.
- STERN, A. The latency of the spiral effect and aftereffect as a function of illumination and speed of rotation. Unpublished doctoral dissertation, University of Connecticut, 1959.
- TAYLOR, M. M. Tracking the decay of the after-effect of seen rotary movement. *Perceptual and Motor Skills*, 1963, 16, 119-129.
- TURNBULL, H. W. The great mathematicians. In J. R. Newman (Ed.), *The world of mathematics*. New York: Simon & Schuster, 1956. P. 147.
- WEYL, H. Symmetry. In J. R. Newman (Ed.), *The world of mathematics*. New York: Simon & Schuster, 1956. Pp. 716-717.
- WOHLGEMUTH, A. On the aftereffect of seen movement. *British Journal of Psychology Monograph Supplements*, 1911, 1, 1-117.

(Received June 3, 1964)

RATING EXTREMITY: PATHOLOGY OR MEANINGFULNESS?

DENIS O'DONOVAN¹

Florida Atlantic University

An attempt is made to reconcile those studies linking the tendency to use the extremes of rating scales (polarization) to pathology with those studies linking polarization to more effective behavior. It is suggested that simultaneous consideration of stimulus and task meaningfulness and subject personality classification will shed light on the seeming discrepancies. A set of propositions consistent with the results of both sets of studies is advanced. Implications of laboratory studies of verbal learning for improved research in response style are discussed.

This article attempts a reconciliation of two contrasting sets of hypotheses about the tendency to use the extremes of rating scales. This tendency is reliable (Peabody, 1962; Rundquist, 1950; Zuckerman, Norton, & Sprague, 1958) and easy to measure, and data have been published in support of each set of hypotheses.

The first set links the use of extreme ratings with pathology. With varying amounts of empirical support, extreme rating style has been suggested as a measure or correlate of neurosis (Johnson, 1946), primitive id impulses (Brenner, 1955), mental-patient status (Borgatta & Glass, 1961), deviance (Berg, 1957), maladjustment (Berg & Collier, 1953; Zax, Gardiner, & Lowy, 1964), anxiety (Lewis & Taylor, 1955), intolerance of ambiguity (Frenkel-Brunswik, 1949; Soueif, 1958), inflexibility (Schutz & Foster, 1963), desire for certainty (Brim & Hoff, 1957), ethnocentrism, authoritarianism, and rigidity (Adorno, Frenkel-Brunswik, Levinson, & Sanford, 1950; Mogar, 1960), and dogmatism (Rokeach, 1960). For instance, Wertheimer and McKinney (1952)

report, "... neurotics exceed the controls in overreaction in using more often the extremes on a rating device [p. 59]."

In contrast, the second set of hypotheses links the use of extreme ratings with meaningful commitment and constructive behavior. Extremeness of ratings has been used or suggested as an operational definition of decisiveness (Cromwell & Caldwell, 1962), saturation of meaning (Mitsos, 1961), committing oneself versus evasiveness (Broen & Wirt, 1958), stimulus importance (Blum, 1964), intensity of meaning (Kanungo & Lambert, 1964), involvement (Saper, 1964), emotional investment (Morris, Eiduson, & O'Donovan, 1960), commitment (O'Donovan, 1960, 1962), and meaningfulness (Isaacson, 1962). These studies have been stimulated by the earlier writings of Morris (1938, 1946, 1956) and Kelly (1955, 1958).

The most emotionally neutral term used for extremity of rating is degree of polarization (Osgood, Suci, & Tannenbaum, 1957), defined operationally as the distance from the point of origin (neutral point) on a rating scale. For convenience, polarization will be used in this article to denote a tendency to rate more extremely; depolarization to

¹ Parts of this paper were completed while the author was at the University of Missouri and at the University of Florida.

denote a tendency to rate closer to the neutral point. While the rating scales used in the studies referred to above differ in various details, in each case the average deviation from the neutral point of the ratings of each subject (or group of subjects) was used to test some hypothesis. In other words, degree of polarization was used as an operational definition of some tendency which was also given a conceptual definition.

Polarization is a fourth and relatively neglected aspect of response style. In his review of response style as a personality variable, McGee (1962) states:

This work has been confined almost exclusively to only three types of response tendency: the *social desirability* set, characterized by the consistent endorsement of desirable traits and the denial of undesirable ones; the *deviation* of a pattern of scores from the typical pattern produced by a given population of responders; and the *acquiescence* set, which consists of tendencies to choose the "true," "agree," or "like" option rather than their respective negative alternatives [p. 284].

The role of polarization in the general study of response set in personality assessment (Berg, 1965; Jackson & Messick, 1962) and its relationships to social desirability (Crowne & Marlowe, 1960; Edwards, 1957), deviation (Berg, 1961), and acquiescence (Couch & Keniston, 1960; Martin, 1964) are not clear at this time.

There is one obvious reason for the neglect of polarization. With the exception of Berg's (1961) four-choice Perceptual Reaction Test (PRT), most of the instruments used in response-style studies have been true-false or dichotomous forced choice in form. Under these limited circumstances, deviation, acquiescence, and social desirability all seem more appropriate for study than polarization. However, the definition of depolarization could be

considered to cover the use of a permissible-evasive response to a true-false item or even the omission of any response. The tendency to avoid committing oneself is well documented as a stable personality characteristic (Cronbach, 1950; Guilford, 1954; Lorge, 1937). Rubin-Rabson (1954) found a negative correlation between this tendency and self-sufficiency. Edwards and Walsh (1964) found this tendency to be unrelated to acquiescence and social desirability. The difficulty with this use of the concept of depolarization is that the instruments provide no opportunity to measure the full range of the polarization-depolarization continuum.

Six studies suggest that polarization (in its more usual definition) is readily distinguishable from acquiescence, as variously defined.

Couch and Keniston (1960) found that response categories *agree* and *disagree* were more highly correlated with overall agreement score (acquiescence) than were *strongly agree* and *strongly disagree*. Schutz and Foster (1963) found one factor for acquiescence and another for polarization. They suggest that the tendency to choose *strongly agree* probably has a different meaning from the tendency to choose *moderately agree*. In a factor analysis of responses to neutral questions, Broen and Wirt (1958) found two nearly orthogonal factors, one of the tendency to agree and the other of the tendency to disagree, rather than one acquiescence factor. Polarization was positively loaded on the first factor. The tendency to be undecided was negatively loaded on both factors.

Zuckerman, Oppenheimer, and Gershowitz (1965) found that actors differed from teachers in polarization but not in acquiescence. Forehand (1962) found polarization measures related to each other, acquiescence measures not

related to each other, and polarization not related to acquiescence. Zuckerman et al. (1958) found acquiescence positively related to authoritarian parental attitudes, and polarization negatively related to these attitudes.

The relationship between polarization and deviation appears clear at first glance: If the typical pattern is polarized, polarization is not deviant, otherwise it is. However, there are a number of complications. Sechrest and Jackson (1962, 1963) point out some, including the problem of the deviantly nondeviant. Johnston (1964) found a group of psychologists who followed the typical cultural value pattern to a greater extent than the typical American, that is, they rated the accepted values higher and the rejected values lower than the norm group. Whether their polarization defines them as deviant, nondeviant, or deviantly nondeviant is not clear.

When social desirability is added, the picture becomes even more clouded. For instance, Zax, Cowen, and Peter (1963) found that novice nuns *deviated* from college females in showing greater *polarization* on ratings of *social desirability*.

While the present discussion is primarily limited to the one "response style" of polarization, its emphasis on meaningfulness places it in fundamental agreement with the recent re-emphasis on item content (McGee, 1965; Peabody, 1964, 1965; Rorer, 1965; Rorer & Goldberg, in press; Samelson, 1964). The present evidence suggests that response style becomes more important than item content as a result of confusion. This confusion could be a result of the pathology or rigidity of the respondent, the ambiguity or meaninglessness of the item (Cronbach, 1950; Gulliksen, 1950), or even the failure of the experimenter to consider relevant stimu-

lus and subject variables. For instance, Zuckerman et al. (1965) found greater polarization for actors than for teachers on the PRT. They suggested two interpretations; "a general behavioral deviancy which expresses itself in an inability to modulate their attitudinal reactions" or "heightened emotionality or drive [p. 170]." The hypothesis that the abstract designs of the test might be more meaningful to the actors was not considered. This is consistent with their statement: "The PRT is the most standardized of the non-content tests [p. 169]." While Berg's (1959) statement of the unimportance of test-item content did not go as far as Zuckerman's just-quoted statement, Berg does reveal that the PRT was designed to be relatively meaningless by Berg, W. A. Hunt, and E. H. Barnes, all of whom are teachers and none of whom are actors.

Meaningfulness is a function of the subject and the stimulus. Pathology in the subject may be expressed in response to any stimulus.

This article will attempt to explore the conditions under which degree of polarization is more predictable from the stereotypic and perhaps pathological response style of the subject, and those conditions under which degree of polarization is more predictable from degree of meaningfulness of the rating task.

POLARIZATION AS PATHOLOGY

Several of the studies previously mentioned found a positive relationship between polarization and such social pathologies as authoritarianism and dogmatism, while Barker (1958), Paine (1964), Peak, Muney, and Clay (1960), and Zuckerman et al. (1958) failed to find such a relationship. The present article deals primarily with individual pathology, that is, the contrast between effective behavior and

neurotic or schizoid behavior. However, one statement about the measurement of social attitudes seems appropriate. When a bigoted person responds to stimuli related to the subject of his bias, both the pathology hypothesis and the meaningfulness hypothesis would predict polarization. Any test of the two hypotheses would necessarily involve unrelated stimuli. An example of the interactions involved is the study by Mogar (1960). He found that polarization increased with more controversial (here translated meaningful) stimuli and with more authoritarian subjects.

Neuringer (1961) hypothesized that a type of thought disorder called the tendency to think in terms of absolute-value dichotomies leads to a higher probability of suicide. Neuringer's operational definition of this tendency was polarization, as defined above. He used the semantic differential (Osgood et al., 1957) with concepts "pregnant with personal meanings" and dimensions heavily weighted on the evaluative factor. He found that suicidal neurotics and psychosomatic patients exceeded general medical and surgical patients (without obvious neurotic disorders) in polarization. Suicidal and psychosomatic patients did not differ from each other. From these results he concluded that: "Dichotomous Evaluative Thinking seems to be a common characteristic of emotionally disturbed persons [p. 449]."

Certain assumptions explicit or implicit in Neuringer's (1961) study can be questioned. He explicitly assumed that the normal hospitalized subjects were under the least stress and did not follow up on the possibility that stress or disinterest might have resulted in constriction of their ratings. He seems to have assumed that the experimental task was equally meaningful for each group of subjects,

though no subjects were consulted on this question. Most important, the statement of the conclusion quoted above seems to include schizophrenics, though no known schizophrenics were included in the study. Zax, Loiselle, and Karras (1960) report no significant differences in extreme ratings between schizophrenics and controls, but greater use by schizophrenics of neutral ratings. It should be added, however, that Zax et al. (1964) found greater use of extremes by schizophrenics.

A more comprehensive reformulation of Neuringer's (1961) conclusion might be that the tendency to respond with similar behavior in dissimilar situations is a common characteristic of disturbed persons. Dichotomous evaluative thinking or polarization, when inappropriate, is simply a special case of this general tendency shown by *some* disturbed persons, notably those with certain cognitive organizations often labeled neurotic.

Long before the current flurry of response-style research, Fenichel (1945) said: "Patients, instead of reacting vividly to actual stimuli, according to their specific nature, react repeatedly with rigid patterns [p. 542]." The three key words in this interpretation are: vividly, suggesting that an intense response is sometimes appropriate; repeatedly, suggesting that more than one sample of behavior is needed to assess rigidity; and patterns, suggesting that no one pattern is *the* rigid pattern.

One example of effective organization is the statement of a scientist quoted by Eiduson (1962): "If you never chase sidelines, you never find anything new; if you chase all the sidelines, you never find anything because you are running down too many blind alleys [p. 126]." This suggests to the writer that an effective person polarizes toward meaningfulness: already

existing meaning in the adjusted person, potential meaning in the creative person. He neither dissipates his energy in intense responses to every trivial stimulus nor misses chances for appropriate polarization.

In contrast, the ineffective person's behavior is more predictable from his customary response style, even when the situation calls for quite different responses. The general case has been called stereotype, rigidity, perseveration, lack of discrimination, etc. Two special cases are polarization and depolarization.

DEPOLARIZATION AS PATHOLOGY

A large number of studies have been conducted on the predictability or stereotypy of behavior in monkeys (Harlow & Zimmerman, 1959; Mason & Green, 1962), chimpanzees (Davenport & Menzel, 1963; Menzel, Davenport, & Rogers, 1963), mental defectives (Berkson & Davenport, 1962), institutionalized infants (Ribble, 1943), schizophrenics (Adams, 1960; Armistage, Brown, & Denny, 1964; Hunter, Schooler, & Spohn, 1962; Luchins, 1959; Painting, 1961; Rodnick & Garmezy, 1957), and process versus reactive schizophrenics (Reisman, 1960; Zlotowski & Bakan, 1963). All of the subjects displaying more stereotypic behavior had been deprived of normal maternal care, by inference in the human subjects, by experiment in the infrahuman subjects. All displayed what is commonly called flattened affect. The evidence of depolarization in these subjects must be considered an analogy. However, anyone observing a mother-raised and a cubicle-raised monkey would probably find it plausible to rate the former as more of a polarizer than the latter.

More direct evidence of depolarization in schizophrenics is found in Bopp (1955), O'Donovan, Morris, and Eidu-

son (1959), and Morris et al. (1960). Evidence of depolarization in less-effective guidance counselors and in schizophrenics is found in O'Donovan (1960).

POLARIZATION AS MEANINGFULNESS

Three studies using the same measure of polarization as Neuringer (1961) illustrate the second set of hypotheses. Mitsos (1961) had subjects choose the 9 (of 21) semantic-differential scales most personally meaningful. Polarization was greater on these scales than on others. Cromwell and Caldwell (1962) found greater polarization on subjects' own dimensions, taken from their Role Construct Repertory Tests (Kelly, 1955), than on dimensions taken from the RCRTs of others. Saper (1964) found greater polarization on a specially designed "involvement" questionnaire for subjects who discussed a case history with each other than for subjects who merely read the case history. Subjects also reported that the group discussion was a meaningful experience.

As can be seen from these examples, the concept meaningfulness has been given various operational definitions. This fact calls for much further careful work but is no bar to continued progress. We can expect that meaningfulness, being a stimulus characteristic, will continue to have more operational definitions than will polarization, a response characteristic. It is very easy to make any stimulus more meaningful; this is done every time an organism is conditioned. The most ambitious attempt at a rigorous hypothetico-deductive system in psychology (Hull, 1943, 1952) limits itself to four basic response measurements while allowing for an infinite variety of stimulus measurements.

Various operations, many of which

are discussed by Underwood and Schulz (1960) and Osgood et al. (1957), have been used to define cultural or consensual meaningfulness.

Most studies using personal or idiographic meaningfulness as an independent variable have followed the reasoning of Allport (1953, 1961) or of Kelly (1955). Allport advises the experimenter to ask the subject. Using this technique, O'Donovan (1964) found greater polarization for value statements checked as meaningful than for those not so checked. Kelly asserts that a person's own dimensions (constructs) are more meaningful to him than are those constructed by other persons. Isaacson (1962) found greater polarization for subjects' own constructs than for those derived from Butler and Haigh (1954) or Osgood et al. (1957).

Informal evidence indicates that subjects differ greatly in what they consider meaningful. They even differ greatly in what they mean by the word "meaningful." Yet they respond readily when asked to rate meaningfulness and proceed as if they know what they are doing. That is, this task seems meaningful, or makes sense, to them.

The foregoing is not meant to predict that the various operational definitions of meaningfulness will be found to refer to the same underlying mechanisms in the individual subject, but to predict that all operational definitions of meaningfulness used in the studies mentioned will be found to correlate positively with polarization.

It is possible to argue that a person becomes the victim of pathological thinking as evidenced by dichotomous evaluation when concepts or scales meaningful to him are used. But it seems equally plausible to speak of meaningfulness leading to strong commitment. A pattern of intense yeasaying and naysaying may be most appro-

priate in the most meaningful situations. Mild affect and tolerance for ambiguity may be out of place when being pursued by a bear or pursuing the girl one loves.

POLARIZATION IN THE LABORATORY

The laboratory situation is here considered a special case of the meaningful situation. This assertion will not be defended here. Several recent discussions (Farber, 1963; O'Donovan, 1963; Orne, 1962; Riecken, 1962) suggest that the subject may find meaning even when none is intended by the experimenter. If it is agreed that the laboratory experience is meaningful and is more meaningful when meaningful stimuli and/or scales are used, then the following relationships between polarization and measures of retention and reliability fit into the general rubric of meaningfulness.

Heretofore neglected in the literature on extreme-response tendencies is evidence relating what looks like polarization to effective learning, psychophysical, and psychometric behavior. A thorough review of these studies would cover virtually the history of psychology, at least from Yerkes and Dodson (1908). Rapaport (1942) summarizes many studies relating intensity of emotional response to retention. Koch (1930) reports a curvilinear relationship between rated pleasantness of examination grades and later recall of these grades, with grades rated "indifferent" being most likely forgotten. Postman and Murphy (1943) found that word pairs either strongly compatible or strongly incompatible with subjects' attitudes were retained better than word pairs only mildly compatible or incompatible.

Pertinent studies focus on meaningfulness and learning. Thorndike (1935) reported superior retention of "cherished" over "worthless" material.

Underwood and Schulz (1960) have thoroughly reviewed relationships between extremity of rating and meaningfulness (particularly pp. 19-25) and between meaningfulness and learning. Noble (1958) deals directly with emotionality or affectivity ratings and meaningfulness. Osgood et al. (1957, pp. 155-159) report a relationship between polarization and speed of response. Earlier, Postman and Zimmerman (1945) had linked intensity of attitude to speed of response. A line of studies dating at least to Henman (1911) reports similar findings. Slamecka (1963) relates speed of response to stimulus meaningfulness. In short, there is evidence linking stimulus meaningfulness, speed of response, and polarization.

It may be that polarization, far from being considered pathological, is now taken for granted in the learning laboratory. For instance, Mechanic (1962) writes: "It may also be assumed that the high-meaningful items will evoke strong differential responses at the outset, while the low-meaningful items will not readily evoke such responses [p. 594]."

What is meaningful when the subject enters the laboratory is related to his prior commitments. Three studies using the Allport-Vernon (1931) or Allport-Vernon-Lindzey (1951) *Study of Values* illustrate this point and its effect on response measures. Havron and Cofer (1957) found easier paired-associate learning for religious subjects when the response word is a religious one than when it is one with politico-economic meaning, and the reverse finding for subjects with strong politico-economic values. Vaughan and Mangan (1963) found greater resistance to group pressure for religious subjects when religious questions are discussed, and for subjects with strong economic values, when economic ques-

tions are discussed. Postman, Bruner, and McGinnies (1948) found that subjects more quickly recognized words related to their stronger values.

Some studies in social psychophysics or person perception, such as those of Sherif and Hovland (1961), directly relate polarization with meaningfulness. Polarization increases with greater involvement of the subject (Hovland & Sherif, 1952; Manis, 1960; Pettigrew, Allport, & Barnett, 1958), with the introduction of rewards or valued stimuli to an ordinary psychophysical-judgment task (Tajfel, 1959; Tajfel & Cawasjee, 1959), and with personal relevance of rating categories (Hastorf, Richardson, & Dornbush, 1958).

Tajfel and Wilkes (1964) maintain that "knowing what is important to the rater may enable us to predict how and when he will tend to use his more extreme judgments [p. 48]." Their conclusion was:

Attributes which appear early and which are repeated frequently in free descriptions of other people tend to be assigned more extreme ratings than attributes which have low frequency and priority and tend to be judged as more important in a person than low ranking attributes [p. 47].

It has long been observed that rating scales and stimuli which do not encourage polarization have poorer reliability. Block (1957, p. 359), studying the phenomenology of emotions, notes that subjects tended to emphasize middle intervals in rating certain affects, such as nostalgia, in contrast to more meaningful or vivid affects, such as love, thus bringing about lower reliabilities. Luria (in Osgood et al., 1957, p. 250) found a correlation of .81 between polarization and test-retest reliability. A direct empirical tie-in of meaningfulness, polarization, and reliability was attempted by Smith and Kendall (1963). In a technique somewhat similar to that of Kelly (1955),

nurses were evaluated in the nurses' own terminology. In contrast to the usual findings with supervisors' ratings, the extremes of these scales were used by raters. Smith and Kendall (1963) conclude that "dimensions meaningful to the raters" were the chief factor in obtaining evaluative rating-scale reliabilities above .97.

COMPARISON OF DIFFERENT STUDIES

There are certain obvious differences between those studies apparently linking polarization with pathology and those apparently linking polarization with meaningfulness. The pathology-polarization hypotheses are supported by studies in which rating dimensions are imposed upon the subject. He has no choice as to which dimensions he uses or which stimuli he rates. We assume that the stimuli and dimensions vary in personal meaningfulness, since when we do ask the subjects, they say that they do. We further assume that motivation for responses of varying polarization is provided, to use Hull-Spence terms, by a mixture of relevant and irrelevant drives and the increase in generalized drive level concomitant with anxiety and neuroticism (Hull, 1952; Janet T. Spence, 1963; K. W. Spence, 1960; Taylor & Spence, 1952).

On the other hand, the meaningfulness-polarization hypotheses are supported by studies in which the subject has an opportunity to provide his own personally meaningful dimensions (Cromwell & Caldwell, 1962; Isaacson, 1962), choose which dimensions or stimuli are more meaningful (Mitsos, 1961; O'Donovan, 1964), or take part in what he or others believe to be a more meaningful experience (Saper, 1964; also laboratory references).

This brings us back to the central question of this discussion: Under what conditions is degree of polarization more predictable from stereotypic and

perhaps pathological response style, and under what conditions is degree of polarization more predictable from meaningfulness of the stimuli being rated and/or the rating scales being used?

Most of the studies cited have measured or categorized only the meaningfulness of the stimuli or the personality classification (e.g., normal versus disturbed) of the subjects. What is needed is a set of propositions leading to predictions of the interaction of stimulus meaningfulness and personality classification. The following propositions are suggested.

BASIC PROPOSITIONS

1. Inappropriate or pathological rigidity in an organism is best studied by measuring lack of differentiation between responses in two or more functionally dissimilar situations, to use Dollard and Miller's (1950) term. In this discussion situations differing in meaningfulness have been stressed. However, this proposition does not depend on any of the other propositions. Even should the present propositions concerning meaningfulness and polarization not prove useful, the reader should still be wary of using one set of responses to functionally similar stimuli as an operational definition of pathological rigidity.

2. The extremeness of an organism's response will depend to some extent on the meaningfulness of the stimuli. If we arbitrarily dichotomize the variable of meaningfulness, this proposition leads to the prediction that response to meaningful stimuli will tend toward the extreme (polarize), while response to meaningless stimuli will tend toward the indifferent (depolarize).

3. The extent to which degree of polarization depends upon the meaningfulness of stimuli is related to other personality characteristics.

4. Effective behavior and lack of emotional disturbance are associated with selective use of extreme responses. The more effective and/or less disturbed the individual, the more probable that Proposition 2 will predict his actual behavior.

5. Ineffective behavior and emotional disturbance are associated with less discriminate use of extreme responses, with less differentiation between meaningful and meaningless stimuli. In the event of total collapse of effective behavior patterns, the predictive power of Proposition 2 will also collapse.

6. As Proposition 2 loses predictive power, propositions based on the individual's usual response style gain predictive power.

7. Predictions can also be based on personality classifications. Two examples follow.

SPECIFIC PROPOSITIONS FOR GROUPS DEFINED BY PATHOLOGY

8. Neurotics and psychosomatic patients polarize both meaningful and meaningless stimuli.

9. Schizophrenics and schizoid persons depolarize both meaningful and meaningless stimuli.

An example of results postdicted by these propositions is found in Lindeman and Adams (1963). Normals rated simple light flashes in a different style than they rated abstract designs, while schizophrenics did not. Their interpretation was that "the smaller number of response differences found in the schizophrenic group would seem to indicate some basic factor common to the schizophrenic process. While the data are not conclusive it is possible that the small number of significantly different responses is the result of poor discrimination or lack of attention frequency found in schizophrenic patients [p. 77]."

Two studies dealing with meaning-

fulness and retention, but not explicitly with polarization, showed results consistent with these propositions. Sherman (1957) discovered that normal prison subjects retained meaningful material—but not nonsense syllables—better than did neurotic subjects. Nidorf (1964) used the serial-anticipation method to test learning of nonsense-syllable lists differing in meaningfulness. Increased meaningfulness facilitated the learning of both normals and schizophrenics, but facilitated the learning of normals significantly more than it did the learning of schizophrenics. Neither of these studies suggested the mechanism responsible for these differences. If we follow the reasoning of the propositions and suggest that Mechanic's (1962) assumption "that the high-meaningful items will evoke strong differential responses at the outset [p. 594]" is at least a parallel of polarization, both of these findings would be expected.

These propositions are also consistent with the results both of studies cited in support of the pathology hypothesis and of those cited in support of the meaningfulness hypothesis. Where meaningfulness is not evaluated, these propositions predict that neurotics will polarize (Neuringer, 1961; Wertheimer & McKinney, 1952), and schizophrenics depolarize (Bopp, 1955; Morris et al., 1960; O'Donovan, 1960). Where meaningfulness is evaluated, meaningful stimuli and/or rating scales lead to polarization (Cromwell & Caldwell, 1962; Isaacson, 1962; Mitsos, 1961; O'Donovan, 1964; Saper, 1964).

For the effective individual (or for the normal, if that is the best approximation available), these propositions predict a response whose vigor correlates with the meaningfulness of the stimuli. For the disturbed or less effective individual, predictions made

from knowledge of the person's usual response style or his personality classification will be more powerful than predictions made from knowledge of the meaningfulness of the stimulus.

IMPLICATIONS FOR FURTHER RESEARCH

More sophisticated techniques of measuring stimulus meaningfulness (Jenkins, Russell, & Suci, 1958; Staats & Staats, 1959), polarization (Blum, 1964; Peabody, 1962), and their relationships (Jenkins, 1960; Koen, 1962; Wimer, 1963) are currently available. For instance, Amster (1964) reports significant effects of contextual pleasantness on polarization and recall. Ziller, Shear, and De Cencio (1964) present evidence of the effect on polarization of the interaction between the status of the rater and the instructions given him. These new techniques and, in general, the findings and more solidly developed methods of the verbal learning and verbal behavior field need to be applied to rating extremity studies. It is clear that whether polarization is considered "sick" or "healthy" has generally depended on whether the particular study focused on stimulus characteristics or subject characteristics. Focusing on the interaction between stimulus and subject characteristics may provide considerable clarity and even a reconciliation of the studies mentioned.

Some important questions in this area may only be answered by studies extending over time. The most certain test of whether a person's cognitive system is rigid or allows change is the *ex post facto* measurement of change. Such studies are now being conducted at the University of Missouri Mental and Hygiene Clinic (Landfield, Nawas, & O'Donovan, 1962). Results to date (Ourth, 1963) conflict with the hypothesis that polarization is a reliable

measure of pathology and suggest that polarizers may be more capable of improvement in psychotherapy than depolarizers. The rating scales used in these studies were derived from the dimensions of personality description used by the individual clients themselves before entering psychotherapy and by their assigned psychotherapists. Following the logic of the propositions, we would expect that those who find their own dimensions and those of their therapists meaningful will polarize, and will later find their psychotherapy more meaningful and hence more successful.

Landfield (1964) had clients in psychotherapy rank in order of "felt usefulness in describing others" the dimensions which they themselves and their therapists had used to describe significant persons. Clients then rated themselves on these dimensions. It was discovered that (a) self-ratings on clients' own dimensions were more polarized, and (b) ratings on dimensions ranked as more useful were more polarized.

Further research based on these propositions might investigate whether successful therapy or other positive experience leads to increased differentiation in degree of polarization to meaningful and meaningless stimuli. As an individual becomes more autonomous or self-actualizing, his responses may more consistently follow Proposition 2. He may also have a clearer, more conscious notion of what is meaningful to him. The ability to state to oneself and to others what is meaningful and to respond accordingly (as predicted by Proposition 2) may emerge as a psychological model of human freedom.

REFERENCES

- ADAMS, H. E. Statistical rigidity in schizophrenic and normal groups measured with auditory and visual stimuli. *Psychological Reports*, 1960, 7, 119-122.

- ADORNO, T. W., FRENKEL-BRUNSWIK, ELSE, LEVINSON, D. J., & SANFORD, R. N. *The authoritarian personality*. New York: Harper, 1950.
- ALLPORT, G. W. The trend in motivational theory. *American Journal of Orthopsychiatry*, 1953, 23, 107-119.
- ALLPORT, G. W. *Pattern and growth in personality*. New York: Holt, Rinehart, & Winston, 1961.
- ALLPORT, G. W., & VERNON, P. E. *A study of values*. Boston: Houghton, 1931.
- ALLPORT, G. W., VERNON, P. E., & LINDZEY, G. *A study of values*. (Rev. ed.) Boston: Houghton-Mifflin, 1951.
- AMSTER, HARRIET. Evaluative judgment and recall in incidental learning. *Journal of Verbal Learning and Verbal Behavior*, 1964, 3, 466-473.
- ARMITAGE, S. G., BROWN, C. R., & DENNY, M. R. Stereotypy of response in schizophrenics. *Journal of Clinical Psychology*, 1964, 20, 225-230.
- BARKER, E. N. *Authoritarianism of the political right, center, and left*. (Doctoral dissertation, Teacher's College, Columbia University) Ann Arbor, Mich.: University Microfilms, 1958, No. 58-2525.
- BERG, I. A. Deviant responses and deviant people: The formulation of the deviation hypothesis. *Journal of Counseling Psychology*, 1957, 4, 154-161.
- BERG, I. A. The unimportance of test item content. In B. M. Bass and I. A. Berg (Eds.), *Objective approaches to personality assessment*. New York: Van Nostrand, 1959. Pp. 83-99.
- BERG, I. A. Measuring deviant behavior by means of deviant response sets. In I. A. Berg & B. M. Bass (Eds.), *Conformity and deviation*. New York: Harper, 1961. Pp. 328-379.
- BERG, I. A. (Chm.) Response set in personality assessment. Symposium presented at Louisiana State University, 1965.
- BERG, I. A., & COLLIER, J. S. Personality and group differences in extreme response sets. *Educational and Psychological Measurement*, 1953, 13, 164-169.
- BERKSON, G., & DAVENPORT, R. K., JR. Stereotyped movements of mental defectives: I. Initial survey. *American Journal of Mental Deficiency*, 1962, 66, 849-852.
- BLOCK, J. Studies in the phenomenology of emotions. *Journal of Abnormal and Social Psychology*, 1957, 54, 358-363.
- BLUM, J. M. *A moving rating scale and the multiple component analysis of evaluation*. (Doctoral dissertation, University of Florida) Ann Arbor, Mich.: University Microfilms, 1964, No. 65-2414.
- BOPP, JOAN. *A quantitative semantic analysis of word association in schizophrenia*. (Doctoral dissertation, University of Illinois) Ann Arbor, Mich.: University Microfilms, 1955, No. 13,458.
- BORGATTA, E. F., & GLASS, D. C. Personality concomitants of extreme response sets (ERS). *Journal of Abnormal and Social Psychology*, 1961, 55, 213-221.
- BRENNER, C. *An elementary textbook of psychoanalysis*. New York: Doubleday, 1955.
- BRIM, O. G., JR., & HOFF, D. B. Individual and situational differences in desire for certainty. *Journal of Abnormal and Social Psychology*, 1957, 54, 225-229.
- BROEN, W. E., JR., & WIRT, R. D. Varieties of response sets. *Journal of Consulting Psychology*, 1958, 22, 237-240.
- BUTLER, J. M., & HAIGH, G. V. Changes in the relation between self-concepts and ideal-concepts. In C. R. Rogers & Rosalind F. Dymond (Eds.), *Psychotherapy and personality change*. Chicago: University Chicago Press, 1954. Pp. 55-75.
- COUCH, A., & KENISTON, K. Yeasayers and naysayers: Agreeing response set as a personality variable. *Journal of Abnormal and Social Psychology*, 1960, 60, 151-174.
- CROMWELL, R. L., & CALDWELL, D. F. A comparison of ratings based on personal constructs of self and others. *Journal of Clinical Psychology*, 1962, 18, 43-46.
- CRONBACH, L. J. Further evidence on response sets and test designs. *Educational and Psychological Measurement*, 1950, 10, 3-31.
- CROWNE, D. P., & MARLOWE, D. A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, 1960, 24, 349-354.
- DAVENPORT, R. K., JR., & MENZEL, E. W., JR. Stereotyped behavior of the infant chimpanzee. *Archives of General Psychiatry*, 1963, 8, 99-104.
- DOLLARD, J., & MILLER, N. E. *Personality and psychotherapy*. New York: McGraw-Hill, 1950.
- EDWARDS, A. L. *The social desirability variable in personality assessment and research*. New York: Dryden, 1957.
- EDWARDS, A. L., & WALSH, J. A. A factor analysis of ? scores. *Journal of Abnor-*

- mal and Social Psychology, 1964, 69, 559-563.
- EIDUSON, BERNICE. *Scientists: Their psychological world*. New York: Basic Books, 1962.
- FARBER, I. E. The things people say to themselves. *American Psychologist*, 1963, 18, 185-197.
- FENICHEL, O. *The psychoanalytic theory of neurosis*. New York: Norton, 1945.
- FOREHAND, G. A. Relationships among response sets and cognitive behaviors. *Educational and Psychological Measurement*, 1962, 22, 287-302.
- FRENKEL-BRUNSWIK, ELSE. Intolerance toward ambiguity as an emotional and perceptual personality variable. *Journal of Personality*, 1949, 18, 108-143.
- GUILFORD, J. P. The validation of an "indecision" score for predicting proficiency of foremen. *Journal of Applied Psychology*, 1954, 38, 224-226.
- GULLIKSEN, H. *Theory of mental tests*. New York: Wiley, 1950.
- HARLOW, H. F., & ZIMMERMAN, R. R. Affective responses in the infant monkey. *Science*, 1959, 130, 421-231.
- HASTORF, A. H., RICHARDSON, S. A., & DORNBUSCH, S. M. The problem of relevance in the study of person perception. In R. Tagiuri & L. Petrullo (Eds.), *Person perception and interpersonal behavior*. Palo Alto: Stanford Univ. Press, 1958. Pp. 54-62.
- HAVRON, M. D., & COFER, C. N. On the learning of material congruent and incongruent with attitudes. *Journal of Social Psychology*, 1957, 46, 91-98.
- HENMON, V. A. The relation of the time of a judgment to its accuracy. *Psychological Review*, 1911, 18, 186-201.
- HOVLAND, C. I., & SHERIF, M. Judgmental phenomena and scales of attitude measurement: Item displacement in Thurstone scales. *Journal of Abnormal and Social Psychology*, 1952, 47, 822-832.
- HULL, C. L. *Principles of behavior*. New York: Appleton-Century-Crofts, 1943.
- HULL, C. L. *A behavior system*. New Haven: Yale Univ. Press, 1952.
- HUNTER, M., SCHOOLER, C., & SPOHN, H. E. The measurement of characteristic patterns of ward behavior in chronic schizophrenics. *Journal of Consulting Psychology*, 1962, 26, 69-73.
- ISAACSON, G. A comparative study of the meaningfulness of personal and cultural constructs. Unpublished master's thesis, University of Missouri, 1962.
- JACKSON, D. N., & MESSICK, S. Response styles and the assessment of psychopathology. In S. Messick & J. Ross (Eds.), *Measurement in personality and cognition*. New York: Wiley, 1962. Pp. 129-155.
- JENKINS, J. J. Degree of polarization and scores on the principal factors for concepts in the semantic atlas. *American Journal of Psychology*, 1960, 73, 274-279.
- JENKINS, J. J., RUSSELL, W. A., & SUCI, G. J. Studies in the role of language in behavior. Technical Report No. 20, 1958, University of Minnesota.
- JOHNSON, W. *People in quandaries*. New York: Harper, 1946.
- JOHNSTON, H. T. *An empirical study of the value orientations of psychologists*. (Doctoral dissertation, University of Missouri) Ann Arbor, Mich.: University Microfilms, 1964, No. 64-13,293.
- KANUNGO, R., & LAMBERT, W. E. Effects of variations in amount of verbal repetition on meaning and paired-associate learning. *Journal of Verbal Learning and Verbal Behavior*, 1964, 3, 358-361.
- KELLY, G. A. *The psychology of personal constructs*. New York: Norton, 1955.
- KELLY, G. A. Man's construction of his alternatives. In G. Lindzey (Ed.), *Assessment of human motives*. New York: Rinehart, 1958. Pp. 33-64.
- KOCH, HELEN L. The influence of some affective factors upon recall. *Journal of General Psychology*, 1930, 4, 171-189.
- KOEN, F. Polarization, *m*, and emotionality in words. *Journal of Verbal Learning and Verbal Behavior*, 1962, 1, 183-187.
- LANDFIELD, A. W. Self conception and meaningfulness as related to own versus therapist's personal construct dimensions. Unpublished manuscript, University of Missouri, 1964.
- LANDFIELD, A. W., NAWAS, M. M., & O'DONOVAN, D. Improvement ratings by external judges and psychotherapists. *Psychological Reports*, 1962, 11, 747-748.
- LEWIS, N. A., & TAYLOR, JANET A. Anxiety and extreme response preferences. *Educational and Psychological Measurement*, 1955, 15, 111-116.
- LINDEMAN, H. H., & ADAMS, H. E. Deviant responses to ambiguous visual stimulus patterns. *Psychological Record*, 1963, 13, 73-77.
- LORGE, I. Gen-like: Halo or reality? *Psychological Bulletin*, 1937, 34, 545-546. (Abstract)
- LUCHINS, A. S. *Rigidity of behavior*. Eugene: Univ. Oregon Press, 1959.

- MANIS, M. The interpretation of opinion statements as a function of recipient attitude. *Journal of Abnormal and Social Psychology*, 1960, 60, 340-344.
- MARTIN, J. Acquiescence—measurement and theory. *The British Journal of Social and Clinical Psychology*, 1964, 3, 216-225.
- MASON, W. A., & GREEN, P. C. The effects of social restriction on the behavior of rhesus monkeys: IV. Responses to a novel environment and to an alien species. *Journal of Comparative and Physiological Psychology*, 1962, 55, 363-368.
- McGEE, R. K. Response style as a personality variable: By what criterion? *Psychological Bulletin*, 1962, 59, 284-295.
- McGEE, R. K. Response set in relation to personality: An orientation. Contribution to I. A. Berg (Chm.), Response set in personality assessment. Symposium presented at Louisiana State University, 1965.
- MECHANIC, A. The distribution of recalled items in simultaneous intentional and incidental learning. *Journal of Experimental Psychology*, 1962, 63, 593-600.
- MENZEL, E. W., JR., DAVENPORT, R. K., JR., & ROGERS, C. M. Effects of environmental restriction upon the chimpanzee's responsiveness in novel situations. *Journal of Comparative and Physiological Psychology*, 1963, 56, 329-334.
- MITSOS, S. B. Personal constructs and the semantic differential. *Journal of Abnormal and Social Psychology*, 1961, 62, 433-434.
- MOGAR, R. E. Three versions of the F Scale and performance on the semantic differential. *Journal of Abnormal and Social Psychology*, 1960, 60, 262-265.
- MORRIS, C. W. Foundations of the theory of signs. *International Encyclopedia of Unified Science*, 1(2). Chicago: Univer. Chicago Press, 1938.
- MORRIS, C. W. *Signs, language, and behavior*. New York: Prentice-Hall, 1946.
- MORRIS, C. W. *Varieties of human value*. Chicago: Univer. Chicago Press, 1956.
- MORRIS, C., EIDUSON, BERNICE, & O'DONOVAN, D. Values in psychiatric patients. *Behavioral Science*, 1960, 5, 297-312.
- NEURINGER, C. Dichotomous evaluations in suicidal individuals. *Journal of Consulting Psychology*, 1961, 25, 445-449.
- NIDORF, L. J. The role of meaningfulness in the serial learning of schizophrenics. *Journal of Clinical Psychology*, 1964, 20, 92.
- NOBLE, C. E. Emotionality (*e*) and meaningfulness (*m*). *Psychological Reports*, 1958, 4, 16.
- O'DONOVAN, D. The value of valuing. In F. McKinney (Chm.), Teaching values. Symposium presented at American Psychological Association, Chicago, September 1960.
- O'DONOVAN, D. Commitment with openness. In S. Cook (Ed.), *Research plans in the fields of religion, morality, and values*. New York: Religious Education Association, 1962. Pp. 67-72.
- O'DONOVAN, D. Questions that can be answered. In H. D. Kimmel (Chm.), Awareness as a factor in verbal operant conditioning. Symposium presented at Southeastern Psychological Association, Miami Beach, April 1963.
- O'DONOVAN, D. Polarization and meaningfulness in 6300 value judgments. Unpublished manuscript, University of Missouri, 1964.
- O'DONOVAN, D., MORRIS, C., & EIDUSON, BERNICE. One way of life in relation to psychological health. *American Psychologist*, 1959, 14, 365. (Abstract)
- ORNE, M. T. On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*, 1962, 17, 776-783.
- OSGOOD, C. E., SUCI, G. J., & TANNENBAUM, P. H. *The measurement of meaning*. Urbana: Univer. Illinois Press, 1957.
- OURTH, L. L. *The relationship of similarity in therapist-client pairs to client's stay and improvement in psychotherapy*. (Doctoral dissertation, University of Missouri) Ann Arbor, Mich.: University Microfilms, 1963, No. 64-1719.
- PAINE, C. B. A tentative model of creativity. Unpublished master's thesis, University of Missouri, 1964.
- PAINTING, D. H. The performance of psychopathic individuals under conditions of positive and negative partial reinforcement. *Journal of Abnormal and Social Psychology*, 1961, 62, 352-355.
- PEABODY, D. Two components in bipolar scales: Direction and extremeness. *Psychological Review*, 1962, 69, 65-73.
- PEABODY, D. Models for estimating content and set components in attitude and personality scales. *Educational and Psychological Measurement*, 1964, 24, 255-269.
- PEABODY, D. Authoritarianism scales and response set. Unpublished manuscript, Swarthmore College, 1965.

- PEAK, HELEN, MUNNEY, BARBARA, & CLAY, MARGARET. Opposites structures, defenses, and attitudes. *Psychological Monographs*, 1960, 74(8, Whole No. 495).
- PETTIGREW, T. F., ALLPORT, G. W., & BARNETT, E. O. Binocular resolution and perception of race in South Africa. *British Journal of Psychology*, 1958, 49, 265-278.
- POSTMAN, L., BRUNER, J. S., & MCGINNIES, E. Personal values as selective factors in perception. *Journal of Abnormal and Social Psychology*, 1948, 43, 142-154.
- POSTMAN, L., & MURPHY, G. The factor of attitude in associative memory. *Journal of Experimental Psychology*, 1943, 33, 228-238.
- POSTMAN, L., & ZIMMERMAN, C. Intensity of attitude as a determinant of decision time. *American Journal of Psychology*, 1945, 58, 510-518.
- RAPAPORT, D. *Emotions and memory*. Baltimore: Williams & Wilkins, 1942.
- REISMAN, J. M. Motivational differences between process and reactive schizophrenics. *Journal of Personality*, 1960, 28, 12-25.
- RIBBLE, MARGARET A. *The rights of infants*. New York: Columbia Univer. Press, 1943.
- RIECKEN, H. W. A program for research on experiments in social psychology. In N. F. Washburne (Ed.), *Decisions, values and groups*. Vol. 2. New York: Pergamon Press, 1962. Pp. 25-41.
- RODNICK, E., & GARMEZY, N. An experimental approach to the study of motivation in schizophrenia. In M. B. Jones (Ed.), *Nebraska Symposium on Motivation: 1957*. Lincoln: Univer. Nebraska Press, 1957. Pp. 109-184.
- ROKEACH, M. *The open and closed mind*. New York: Basic Books, 1960.
- RORER, L. G. The great response-style myth. *Psychological Bulletin*, 1965, 63, 129-156.
- RORER, L. G., & GOLDBERG, L. R. Acquiescence in the MMPI? *Educational and Psychological Measurement*, in press.
- RUBIN-RABSON, G. Correlates of the non-committal test-item response. *Journal of Clinical Psychology*, 1954, 10, 93-95.
- RUNDQUIST, E. A. Response sets: A note on consistency in taking extreme positions. *Educational and Psychological Measurement*, 1950, 10, 97-99.
- SAMELSON, F. Agreement set and anti-content attitudes in the F Scale. *Journal of Abnormal and Social Psychology*, 1964, 68, 338-342.
- SAPER, M. B. Involvement in reading a case history. Unpublished master's thesis, University of Missouri, 1964.
- SCHUTZ, R. E., & FOSTER, R. J. A factor analytic study of acquiescent and extreme response set. *Educational and Psychological Measurement*, 1963, 23, 435-447.
- SECHREST, L., & JACKSON, D. N. The generality of deviant response tendencies. *Journal of Consulting Psychology*, 1962, 26, 395-401.
- SECHREST, L., & JACKSON, D. N. Deviant response tendencies: Their measurement and interpretation. *Educational and Psychological Measurement*, 1963, 23, 33-53.
- SHERIF, M., & HOVLAND, C. I. *Social judgment: Assimilation and contrast effects in communication and attitude change*. New Haven: Yale Univer. Press, 1961.
- SHERMAN, L. J. Retention in psychopathic, neurotic, and normal subjects. *Journal of Personality*, 1957, 25, 721-729.
- SLAMECKA, N. J. Choice reaction-time as a function of meaningful similarity. *American Journal of Psychology*, 1963, 76, 274-280.
- SMITH, PATRICIA, & KENDALL, L. M. Re-translation of expectations: An approach to the construction of unambiguous anchors for rating scales. *Journal of Applied Psychology*, 1963, 47, 149-155.
- SOUEIF, M. I. Extreme response sets as a measure of intolerance of ambiguity. *British Journal of Psychology*, 1958, 49, 329-334.
- SPENCE, JANET T. Learning theory and personality. In J. M. Wepman & R. W. Heine (Eds.), *Concepts of personality*. Chicago: Aldine, 1963. Pp. 3-30.
- SPENCE, K. W. *Behavior theory and learning*. Englewood Cliffs, N. J.: Prentice-Hall, 1960.
- STAATS, A. W., & STAATS, CAROLYN K. Meaning and *m*: Correlated but separate. *Psychological Review*, 1959, 66, 136-144.
- TAJFEL, H. The anchoring effects of value in a scale of judgments. *British Journal of Psychology*, 1959, 50, 294-304.
- TAJFEL, H., & CAWASJEE, S. D. Value and the accentuation of judged differences: A confirmation. *Journal of Abnormal and Social Psychology*, 1959, 59, 436-39.
- TAJFEL, H., & WILKES, A. L. Salience of attributes and commitment to extreme judgments in the perception of people.

- British Journal of Social and Clinical Psychology*, 1964, 3, 40-49.
- TAYLOR, JANET A., & SPENCE, K. W. The relationship of anxiety level to performance in serial learning. *Journal of Experimental Psychology*, 1952, 44, 61-64.
- THORNDIKE, E. L. *The psychology of wants, interests, and attitudes*. New York: Appleton-Century, 1935.
- UNDERWOOD, B. J., & SCHULZ, R. W. *Meaningfulness and verbal learning*. New York: Lippincott, 1960.
- VAUGHAN, G. M., & MANGAN, G. L. Conformity to group pressure in relation to the value of the task material. *Journal of Abnormal and Social Psychology*, 1963, 66, 179-183.
- WERTHEIMER, RITA, & MCKINNEY, F. A case history blank as a projective technique. *Journal of Consulting Psychology*, 1952, 16, 49-60.
- WIMER, CYNTHIA. An analysis of semantic stimulus factors in paired-associate learning. *Journal of Verbal Learning and Verbal Behavior*, 1963, 1, 397-407.
- YERKES, R. M., & DODSON, J. D. The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology*, 1908, 18, 459-482.
- ZAX, M., COWEN, E. L., & PETER, MARY. A comparative study of novice nuns and college females using the response set approach. *Journal of Abnormal and Social Psychology*, 1963, 66, 369-375.
- ZAX, M., GARDINER, D. H., & LOWY, D. G. Extreme response tendency as a function of emotional adjustment. *Journal of Abnormal and Social Psychology*, 1964, 69, 654-657.
- ZAX, M., LOISELLE, R. H., & KARRAS, A. Stimulus characteristics of Rorschach ink blots as perceived by a schizophrenic sample. *Journal of Projective Techniques*, 1960, 24, 439-443.
- ZILLER, R. C., SHEAR, H. J., & DECENCIO, D. A professional response set—dogmatism. *Journal of Clinical Psychology*, 1964, 20, 299-303.
- ZLOTOWSKI, M., & BAKAN, P. Behavioral variability of process and reactive schizophrenics in a binary guessing task. *Journal of Abnormal and Social Psychology*, 1963, 66, 185-187.
- ZUCKERMAN, M., NORTON, J., & SPRAGUE, D. S. Acquiescence and extreme sets and their role in tests of authoritarianism and parental attitudes. *Psychiatric Research Reports*, 1958, No. 10, 28-45.
- ZUCKERMAN, M., OPPENHEIMER, CYNTHIA, & GERSHOWITZ, D. Acquiescence and extreme response sets of actors and teachers. *Psychological Reports*, 1965, 16, 168-170.

(Received June 8, 1964)

INFORMATION ABOUT SPATIAL LOCATION BASED ON KNOWLEDGE ABOUT EFFERENCE¹

LEON FESTINGER AND LANCE KIRKPATRICK CANON

Stanford University

An experiment was designed to determine whether or not the human organism possessed "outflow" information derived from monitoring nerve impulses in motor pathways. The experiment focused on the extraocular muscles since proprioceptive input to the central nervous system from these muscles is poor. The results show that in the absence of good proprioceptive information, the presence or absence of "outflow" information makes a difference in accuracy of localizing an object in space.

The human being continually acquires and uses information about himself and his relation to the environment. We are accustomed to thinking of this information as having been acquired through input to afferent mechanisms. That is, we know about the environment through seeing, hearing, touching, and a variety of other means. Not the least of these sources of information is input from proprioceptors. For example, if I am led blindfolded into a room and I touch an object in that room with my hand, I know where that object is in relation to my body because, on the basis of proprioceptive feedback, I know where my hand is.

There is, however, another possible source of information about one's relation to the environment that has not been adequately explored. If, in the central nervous system, outgoing motor nerve impulses are monitored and recorded, then information would also exist concerning spatial location on the basis of this record of efferent impulses,

that is, a record of the specific directions given to the musculature. This information, if it exists, need not rely on any current afferent input. To make this clear, let us illustrate by a loose analogy. Imagine there is a person who will unconditionally obey your orders. Let us also assume that you and the other person have had sufficient previous experience with the environment so that you can give him, and he can follow, clear directions. You tell this person to go to a certain specific place and to wait there for you. Even in the complete absence of any *current* sensory input you will know exactly where that person is because you know where you told him to go.

The question of whether or not such monitored efferent information exists is an old one in psychology although, of late, it has been rarely mentioned. Actually, a closely related speculation was vigorously debated many years ago. James (1950)² stated the issue clearly:

There must, of course, be a special current of energy going out from the brain into the

¹ This research was supported by Grant No. MH 07835-01 from the National Institutes of Health to the senior author. We wish to thank Douglas H. Lawrence and Gordon H. Bower for their help on the experiment.

² We give the dates of the later editions from which we have quoted. The book by James was originally published in 1890, and the first edition of the book by Helmholtz was earlier than that.

appropriate muscles during the act; and this outgoing current (it is supposed) must have in each particular case a feeling *sui generis* attached to it, This feeling of the current of outgoing energy has received from Wundt the name of the *feeling of innervation*. I disbelieve in its existence, and must proceed to criticise the notion of it, at what I fear may to some prove tedious length [p. 493].

If in this statement we replace the phrase "feeling of" by "information about," then this old controversy is exactly germane to our present question. While we do not intend to engage in an exhaustive review of the argument about "feeling of innervation," let us look at the principal data about which the disagreement centered.

One major piece of evidence at that time is summarized by Helmholtz (1925). He states:

For instance, if the external rectus of the right eye is paralyzed or the nerve leading to it, this eye can no longer be pulled around to the right. As long as the patient continues to turn it inwards only it still makes regular movements, and he perceives correctly the directions of objects in the field of view. But the moment he tries to turn his eye outwards, that is, to the right, it ceases to do his bidding, and remains standing in the middle, while the objects appear to move to the right, although the adjustment of the eye and the positions of the retinal images in it have not varied [p. 245].

From this Helmholtz concludes that since there was absolutely no afferent change when the eye tried to move to the right, and since motion was perceived as if the eye *had* moved to the right with the retinal image remaining constant, there must be a feeling of (information about) innervation.

William James (1950) quotes other data in addition. He says:

Partial paralysis of the same muscle, paresis, as it has been called, seems to point even more conclusively to the same inference, that the will to innervate is felt independently of

all its afferent results. I will quote the account given by a recent authority, of the effects of this accident: "When the nerve going to an eye muscle, e.g., the external rectus of one side, falls into a state of paresis, the first result is that the same volitional stimulus, which under normal circumstances would have perhaps rotated the eye to its extreme position outwards, now is competent to effect only a moderate outwards rotation, say of 20 degrees. If now, shutting the sound eye, the patient looks at an object situated just so far outwards from the paretic eye that this latter must turn 20 degrees in order to see it distinctly, the patient will feel as if he had moved it not only 20 degrees toward the side, but into its extreme lateral position, The test proposed by von Graefe [1878], of localization by the sense of touch, serves to render evident the error which the patient now makes. If we direct him to touch rapidly the object looked at, with the fore-finger of the hand of the same side, the line through which the finger moves will not be the line of sight directed 20 degrees outward, but will approach more nearly to the extreme possible outward line of vision [p. 507]."

The theoretical relevance of this observation is stated succinctly by James:

It appears as if here the judgment of direction *could* only arise from the excessive innervation of the rectus when the object is looked at. All the afferent feelings must be identical with those experienced when the eye is sound and the judgment is correct. The eyeball is rotated just 20 degrees in the one case as in the other, the image falls on the same part of the retina, the pressures on the eyeball and the tensions of the skin and conjunctiva are identical. There is only one feeling that *can* vary, and lead us to our mistake. That feeling must be the effort which the will makes, moderate in one case, excessive in the other, but in both cases an efferent feeling, pure and simple [p. 508].

James then proceeds to rebut the interpretations of these observations. Acknowledging that G. E. Müller was the first to propose the rebuttal explanation, he states:

Beautiful and clear as this reasoning seems to be, it is based on an incomplete inventory of the afferent data. The writers have all omitted to consider what is going on in the

other eye. This is kept covered during the experiments, to prevent double images, and other complications. But if its condition under these circumstances be examined, it will be found to present certain changes which must result in strong afferent feelings. And the taking account of these feelings demolishes in an instant all the conclusions which the authors from whom I have quoted base upon their supposed absence [p. 508].

James then proceeds to point out that the covered, healthy eye does rotate as directed by the efferent impulses and thereby provides the *afferent* stimulation necessary for the perception of motion in the Helmholtz (1925) example, or the misperception of direction in the Graefe (1878) example. Although James, in his explanation, never copes with the question of why the afferent impulses from the covered eye should completely dominate the afferent impulses from the open eye, nevertheless his "demolition" of the argument for feeling of (information about) innervation appears to have been very effective. So persuasive was the argument by James that Mach (1914), who in 1886 had argued strongly for the "feeling of innervation" and presented original experiments supporting it, almost completely reversed his stand in the fifth edition of his book, written in 1906. Here he says:

The theory of James and Münsterberg fits these facts, as I think, without any straining, and we ought therefore to consider it as correct in essentials. The innervation is not felt, but the consequences of the innervation set up new peripheral sensible stimuli, which are connected with the execution of the movement [p. 176].

Rightly or wrongly, James apparently won the argument, and the issue has been a dead one in psychology for many years. Many dead issues do not stay dead, however, and this one has recently been revived by physiologists. Recently von Holst (1954), concern-

ing himself with how the organism differentiates between self-generated movement of a part of the body and an identical movement generated by external forces, proposed the idea of "efference copy." His idea was that incoming afferent signals were matched against a temporary copy of outgoing efferent signals. If they matched perfectly, the motion involved was entirely self-generated. This, of course, is somewhat different from the idea that information from a record of efferent impulses is available and used all by itself. Nevertheless, it is related and served to revive the issue in other contexts.

The question has become particularly important to those who are concerned with understanding the control system for eye movements. Probably the major reason for this is that there is great doubt among physiologists that afferent signals from the extraocular muscles are used to any significant extent in the control of eye movements. If afferent feedback from the extraocular muscles is not useful for determining the position of the eye, then it becomes convenient for the theoretician to posit the existence of information obtained from a record of efferent impulses.

Thus, Fender (1964), discussing the possible role of afferent signals of position of the eye, says:

There is experimental evidence that the positioned signal is not used, for if a subject views two similar but separately generated stabilized images, one with each eye, it is found that for a short period the two visual axes move in conjunction. However, this motion quickly breaks down, and the visual axes move independently, sometimes getting as far apart as 30 deg in the horizontal direction and 15 deg in the vertical. There is, of course, no binocular retinal-image disparity to act as a cue in this case, and it appears that any positional signal which might arise from the extraocular muscles is quite ineffective in maintaining the parallelism of the visual axes [p. 317].

Fender proceeds to incorporate an "efferent copy" feedback loop into his model of the physiological system controlling eye movements.

Whitteridge (1962) recently summarized the problem as follows:

The role of extraocular afferent impulses in perception is very uncertain. It is self evident that we are not directly aware of the position of our eyes in the same sense in which we are aware of the position of our fingers even with the eyes shut. The question is whether the position of the eyes enters into judgments of position and movement, and if it does, how far proprioceptors are responsible. The alternative theories are that information from the volume of outgoing motor nerve impulses in the oculomotor pathways is centrally available—the *outflow* theory, or that impulses from proprioceptors directly signal the state of the eye muscles—the *inflow* theory. The strongest point against the inflow theory is that when a patient with a paralyzed and therefore immobile eye tries to turn it to one side, the observed visual field moves as though he had succeeded in moving the eye. This cannot be due to any conceivable change in proprioceptive discharge [p. 511].

As of 1962, among physiologists, the entire controversy seems to have revived. The issue is now more sophisticated from a theoretical point of view; but on the empirical side, Whitteridge (1962) seems to be back to Helmholtz (1925). There is, however, more empirical evidence on the issue today than there was 60 to 70 years ago. Brindley and Merton (1960) report a very direct attempt to settle the question as to whether or not there is usable proprioceptive feedback from the extraocular muscles. They anesthetized the surface of the eyes and the inner surface of the eyelids of subjects and covered the corneas with opaque caps so that the subjects received no visual information. They then mechanically moved a subject's eyeball by catching hold of the insertion of either the medial or lateral rectus muscle with toothed forceps. When the eye was

moved in this manner through rotations of 20 degrees or more, sometimes even backward and forward quite rapidly, the subject did not know that his eye was moving.

Cognizant of the argument offered by James, they repeated these observations moving both eyes simultaneously and obtained the same result. Merton's (1964) paper, the main purpose of which is "... to reinstate the experiments of Helmholtz, which proved that no information about the position of the eyes is derived from sense endings in the eye muscles [p. 315]," comes to the conclusion: "A subject is only conscious of his intention to move his eye and does not know whether the movement has in fact taken place or not [p. 318]."

Considering these new data, it seems highly likely that Helmholtz (1925) was correct and that James (1950), in spite of having won the argument in his day, was wrong. It would be useful, however, to have additional data on the question. After all, the work of Brindley and Merton (1960) demonstrates the absence of a position sense in the eye based solely on proprioception from the extraocular muscles. To strengthen the argument one might well desire positive evidence that information obtained from a record of outgoing motor nerve impulses is available and useful.

Let us be specific. If it is true, as seems likely, that we know the position of the eye mainly in terms of knowing where the eye was directed to go, then it should be possible to show that when the eye is directed to go to a specific location, a subject knows where his eye is more accurately than if the eye arrived at the same position without directions concerning this specific location ever having been issued.

The technical problem in doing such an experiment is, of course, the prob-

lem of how to devise a method of having the subject move his eyes to a specific location without issuing efferent signals concerning that location. A plausible solution to this technical problem may be found in the work of Rashbass (1961). He reports a series of experiments designed to elucidate the relationship between the usual saccadic eye movements and the smooth eye movements that occur in tracking a target. Several of his findings are important to us here.

Rashbass reports evidence that saccadic eye movements and smooth tracking eye movements are controlled and generated independently of one another. Barbiturate drugs serve to almost completely disrupt smooth eye-tracking movements but do not interfere with precise saccadic movements. Thus, a subject who watched a target which moved horizontally at a rate of 3.5 degrees per second ordinarily showed a smooth eye movement before the administration of any drug. After administration of a barbiturate, Rashbass (1961) states:

The first noticeable effect was the increase in the number of saccadic movements occurring during the first second of tracking. As the amount of drug given increased, the saccadic movements increased at the expense of the smooth movements, until, after 8 minutes, no smooth tracking movement could be detected [pp. 333-334].

From this and other data, he concludes that barbiturate drugs make the smooth tracking response inoperative but do not interfere with accurate saccadic eye movements. Hence the two types of eye movements must be separately controlled.

Rashbass also reports data from experiments designed to discover what produces smooth and saccadic eye movements. The specific question is "whether smooth movements are brought about by the position of the

target's image on the retina, or by its movement over the retina [p. 331]." He tests this "by imparting to an initially stationary target a displacement to one side, and at the same time beginning a movement of uniform velocity toward the opposite side [p. 331]." The result is stated by Rashbass as follows:

. . . after a reaction time during which the eye does not move, a smooth movement starts in the direction in which the target is moving. When this has been established, a saccadic movement occurs in the direction opposite to the smooth movement to counteract the lead which the eye has over the target. . . . This result indicates that the smooth movement is stimulated by the movement of the target irrespective of its position. The conclusion that the smooth movements are brought about by the movement of the target explains the apparently paradoxical observation that the first movement which the eye makes may take the point of fixation further from the target than if no eye movement at all were to occur [p. 332].

From this and other data, the conclusion is that "the smooth movement is stimulated by the direction of movement and the velocity of the target, and the saccadic movement is stimulated independently by the position of the target [p. 333]."

We have dealt at length with the results obtained by Rashbass because they are critical for us. They suggest that if the eye were brought into a given position by a saccadic movement, this movement would be a response to efferent signals concerning the position of the target. If, however, the eye were brought into that same position by a smooth tracking movement, the efferent directions would be concerned with velocity matching and not precisely with target location.

A possible experiment suggests itself to answer the question concerning the availability of information based on efference. The experiment would be conducted in a completely dark room

with the subject's head fixed so that only eye movements could occur. In one variation, a light would suddenly appear within the visual field of the subject, then disappear, and the subject would be asked to point to the location where the light had been. In this variation, in order to fixate the light, a saccadic eye movement would occur, and "directions" would have been given to the extraocular muscles to move from "normal frontal" position to a specific location. If these directions to the musculature are monitored and recorded so as to be available as information, the person would know the location of the light on the basis of knowing where he had sent his musculature in order to fixate the light.

In another experimental variation, the light would appear and move slowly and smoothly across the visual field before coming to a stop. The subject would fixate the light when it first appeared and would then track the light across the visual field until it stopped moving. To the extent that *only* smooth tracking eye movements occurred, the musculature would, presumably, simply have been directed to "follow the light." Thus, in this experimental variation, the efferent information that existed would contain information about the direction of movement and the velocity of movement, but would not include information concerning the specifically designated position in which the light had stopped.

In both of the above variations, of course, there would be the same amount of proprioceptive information concerning where the light was. If the subject's head is clamped so that only eye movements are used to fixate the target, the only proprioceptive signals would come from the extraocular muscles. Since these signals are not useful, as Merton (1964) has shown,

then subjects would know the location of the light more accurately when it suddenly appeared than when they tracked it across the visual field. We would, of course, expect more than zero knowledge of location in the tracking condition. The subject would have knowledge about direction and also some knowledge of eye position from afference from the eyelids. Also, it is well known that smooth tracking movements lag and saccadic movements occur periodically. These would also provide additional information. If, however, information based on efference is available, we would expect a difference between the two conditions.

Along with this, of course, one would want to set up another experimental condition in which the subject's head was not clamped so that head movements could be employed in helping to fixate the light. Useful proprioceptive input would be expected from the neck muscles, and to the extent that the position of the light could be adequately known on the basis of these proprioceptive signals from the neck muscles, the difference between the two experimental variations would be expected to vanish.

Such an experimental design, using two manners of presentation of the light and two degrees of adequacy of proprioceptive information, should provide data that would answer the question as to whether or not outflow information is available and is used.

PROCEDURE

Twenty-eight college students, 12 female and 16 male, were subjects in the experiment. Each subject volunteered and was paid \$1.50 for participating.

The experiment was conducted in a light-proof room. The apparatus consisted of an overhead boom fastened to the ceiling with its pivot point slightly in front of a point directly over the subject's head. The boom extended 4 feet forward from the pivot point.

From the far end of the boom hung an illuminated rectangle that measured 2×3 inches. The experimenter, standing to the side of the seated subject, could move the boom noiselessly so that the light was at any desired lateral position. The height of the light was fixed approximately at the subject's eye level. Calibration at the pivot point of the boom enabled the experimenter to read the setting of the light in angular deviation from straight ahead of the subject.

On a table directly in front of the subject and at a suitable height was a pointer attached to a calibrated turntable. The pivot of the pointer was directly underneath the pivot of the boom. The subject, when pointing to where the light was, or had been, was instructed to lay his index finger along the pointer and move it so that he pointed in the proper direction. The measuring scales for both the boom and the pointer were very dimly illuminated and shielded from the subject. The illumination was sufficient, however, to allow the experimenter to read the scales in an otherwise totally dark room. The target light was also dimly illuminated so that there were no problems with after-images, and the target light did not make other things in the room visible.

Fourteen of the subjects, seven male and seven female, were used in the "eye-movement-only" condition. These subjects had their head in a rigid clamp throughout the experiment so that fixating and tracking the target light could be done only with eye movement. The head and body were always in the directly forward position. The other 14 subjects, 9 male and 5 female, were used in the "head-movement" condition. This condition was identical to the other except that the head was not clamped. Thus, these subjects could and did rotate their heads, and even their bodies to some extent, in addition to moving their eyes in fixating and tracking the target light.

Before data collection started, each subject was given practice at using the pointer with the target light at various positions. This practice was continued until the subject was familiar with the situation and the use of the pointer. The actual data collection consisted of 28 trials, 4 trials at each of 7 positions of the light. The positions used were +30, +20, +10, 0, -10, -20, and -30 degrees (+ referring to positions to the subject's right, -, to positions to the subject's left). For one trial in each position the target light was turned on in that position and stayed on. The subject pointed to the light while it was still visible. This

was intended to yield a measure of the accuracy to be expected with optimal information. For another trial in each of the seven positions the light was turned on in that position, stayed on for 3 seconds, and was then turned off. The subject was asked to point, after the light was turned off, to where the light had been. In this condition, outflow information would presumably be available to the subject. When the light came on, the subject would have to direct a saccadic movement of his eyes to a specific location and would, hence, know this location at least with respect to a normal frontal reference point.

On the two remaining trials at each of the target-light positions, the light moved across part of the visual field. The light would appear, move slowly (approximately 10 degrees per second) across the visual field through an angle of 15, 20, 25, 30, or 35 degrees, and come to a halt at the desired position. The light then remained on in this final position for 3 seconds and was then turned off. The subject was asked to point to where the light had been after it was turned off. For each of the seven positions the light moved from right to left on one trial and from left to right on the other trials. These trials were, of course, intended to be trials on which outflow information concerning target position would be less available to the subject. To the extent that smooth tracking eye movements would have been involved, directions concerning target location would not have occurred.

The decision to keep the light on its final position for 3 seconds before turning it off was an arbitrary one. We wanted a period of time long enough so that in the tracking trials there would be no ambiguity about when and where the light had come to a stop. On the other hand, we wanted the period short enough so as to reduce the likelihood of blinking or moving the eyes to a forward position and refixating the light. Such eye movements would tend to vitiate the procedure. Certainly, in 3 seconds such movements can occur, but some compromise between allowing this and having an unambiguous final position was necessary.

The order of trials was arranged in a sequence so that the target light was never in the same position on any two consecutive trials, and the four different kinds of trials were distributed evenly through the series. The same order was used for all subjects. After the subject had pointed for a trial, he was asked to return his hand to his lap. The experimenter then recorded the setting of

the pointer and moved the boom to the appropriate position for the next trial. The interval between trials was approximately 45 seconds. The subject's hand remained in his lap until the experimenter said, "All right, now point to where the light is (was)."

RESULTS

We are interested in the magnitude of the error made by the subject in pointing to the position of the target light. The less adequately the person knows the position of his eyes, or his head, when fixating the light, the less accurate should he be in pointing to its location afterwards. The simplest calculation is, of course, to take the absolute deviation of the pointer position from the target position for each trial. Thus, if the target was in position +20 and the subject set his pointer to +16, this would be an error of 4 degrees. Table 1 presents the results from the experiment based on this simple calculation.

Even a cursory look at the data in Table 1 reveals that the obtained data are of the kind one would expect if, indeed, proprioceptive input from the extraocular muscles is poor and the person has available, and uses, outflow information. When only eye movements are allowed, that is, when the head was clamped, the error of point-

ing when the light suddenly appeared at the designated position was only slightly worse than when the light was on while the person was pointing. However, when the subject tracked the light across the visual field, and thus would not have relevant outflow information, the error of pointing is considerably greater. It is also clear that when head movements are allowed, the results are very different. The "tracking" trials are then slightly superior to the "at position" trials.

We have presented these data because some readers might consider this the proper measure to use. We will not engage in extended discussion of Table 1, however, nor present statistical analyses, since it seems to us that a more accurate measure should be used. The absolute error of pointing is, of course, affected by constant errors. One subject may consistently point somewhat to the right of the target, another consistently to the left. Such constant errors are probably due to coordinating the physical act of pointing with knowledge of location and probably should be disregarded in our calculations. Actually, there was an average constant error of pointing somewhat to the left of the position of the target. Over all types of trials, this average constant error was 1.6 degrees to the left in the "eye-movement-only" condition and .1 degree to the left in the "head-movement" condition. The probable reason for the direction of the constant error in the "eye-movement-only" condition is that, using the right hand, the hand position was more comfortable along the fixed pointer when pointing toward the left than when pointing toward the right. Apparently, head movements provided enough additional orientation to eliminate this constant error.

There is also another source of constant error in the data. Two types of

TABLE 1
AVERAGE ABSOLUTE ERROR (IN DEGREES)
OF POINTING TO TARGET LIGHT

Condition	Type of trial			
	Light on	Light off		
		At position	Tracked from right	Tracked from left
Eye movement only	3.06	3.54	5.24	5.55
Head movement	2.13	3.92	3.69	3.35

tracking trials, one from the left, one from the right, were used because of the possibility that the memory of where the light had stopped might be affected by the direction in which the light had moved. This, indeed, turns out to be the case. In the tracking trials, the subjects tend to point a bit more in the direction from which the light had come. Thus, in the "eye-movement-only" condition the constant error is -1.2 degrees when the light came from the right, but -2.2 degrees when the light moved from the left. Similarly, in the "head-movement" condition the corresponding constant errors are $+1.1$ and $-.6$. The difference between the two types of tracking trials is not quite significant statistically for the "eye-movement-only" condition ($t = 1.44$) but is significant at the 2% level for the "head-movement" condition ($t = 2.85$).

Clearly, we do not want to have our measure of accuracy of pointing contaminated by these various sources of constant error. We, therefore, computed a "corrected absolute error" of pointing by taking into account for each subject, for each type of trial, the constant error in the data. Thus, for example, a subject may have had a

constant error of 2 degrees to the left on the seven trials on which the target was tracked in from the left. If this subject set his pointer at -24 degrees when the target light had actually stopped at -20 degrees, his corrected absolute error on this trial was 2 degrees. Table 2 presents the data using this measure.

These corrected data show the same overall pattern of results as the data using the uncorrected absolute error. We will discuss these data in detail, presenting appropriate statistical analyses.

Eye-Movement-Only Condition

It is clear that when only eye movements are permitted, localization of the target light is better when the light suddenly *appears* at its final position than when it is *tracked* to its final position. An analysis of variance yields a highly significant F value (8.71 , $df = 3/39$) for the variance among the means of the different types of trials. The variance among subjects is also significant ($F = 2.68$, $df = 13/39$). This latter, of course, simply means that some subjects are consistently more accurate than others in pointing to the target light.

The difference in accuracy between the "light-on" and "light-off-at-position" trials is not significant ($t = 1.10$). The mean for each is, however, significantly different from the mean for each of the "tracking" trials, the smallest t value being 3.86 between the "at-position" mean and the "tracked-from-right" mean. In short, with only eye movements permitted, pointing to the target when it suddenly appeared at its final position is not materially less accurate than when the pointing was done while the light was still on. In the tracking conditions, however, when relevant outflow information was presumably not avail-

TABLE 2

AVERAGE CORRECTED ABSOLUTE ERROR (IN DEGREES) OF POINTING TO TARGET LIGHT

Condition	Type of trial			
	Light on	Light off		
		At position	Tracked from right	Tracked from left
Eye movement only	2.58	3.11	4.38	4.50
Head movement	1.95	3.64	3.11	2.79

able, accuracy is materially and significantly worse.

Head-Movement Condition

When head movements are allowed, the data present quite a different pattern, although significant differences still exist among the different types of trials. The variance of the means for the different types of trials and of the means for subjects both yield highly significant F values (7.30, $df = 3/39$; and 6.56, $df = 13/39$).

With head movements, the accuracy of pointing with the light still on is significantly better than each of the three conditions in which the pointing was done after the light was off. The important differences to us, however, are between the "at-position" trials and the "tracking" trials. Here we find that the "at-position" accuracy is no longer better, but is actually worse than the accuracy of pointing on the "tracking" trials. The two t values are 2.01 and 2.19 which, for $df = 13$, are each significant at about the 5% level. We had not anticipated this, and we are not certain of the reason for it. It may simply be that occasional inattention affected accuracy in the "at-position" trials. There was no warning of when the light would appear. In the "tracking" trials, the period of tracking could minimize the effects of any inattention. It is clear, however, that when head movements are allowed, thus making available good proprioceptive input concerning position, the availability of relevant outflow information no longer produces greater accuracy.

Comparison of the Two Conditions

If we compare the accuracy between the condition in which only eye movements were allowed and the condition in which head movements were also

allowed, we see that in the latter condition there is a general tendency to be more accurate. When the light is on while pointing, the average corrected error decreases from 2.58 to 1.95, a difference significant at the 10% level ($t = 1.73$, $df = 26$). The data for the "tracking" trials also show much less error with head movement allowed. The two t values here are 2.11 for "tracking from the right" and 2.98 for "tracking from the left," significant at the 5% and 1% level respectively.

Only for the "light-off-at-position" trials is there no improvement from the "eye-movement-only" to the "head-movement" condition. The actual difference is slightly in the opposite direction but is negligible ($t = .81$). Indeed, it seems as though the presence of relevant outflow information about eye position in the "eye-movement-only" condition is just as good as the presence of the same outflow information plus good proprioceptive input in the "head-movement" condition. It is clear also that, when there is good proprioceptive input and no relevant outflow information, as in the tracking trials with head movements, accuracy is at least as good as when relevant outflow information is also present. This would tend to imply that, in this situation, there is some redundancy of information.

DISCUSSION

The main conclusion we would like to draw from the results of the experiment is that information based on some kind of record of efferent impulses (i.e., outflow information) is available to the organism. The major result on which we wish to base this conclusion is the finding that, when only eye movements were permitted, target localization was more accurate when the target suddenly appeared at

its final position than when it was tracked to that position.

Let us review the line of reasoning involved in coming to this conclusion.

Accuracy of localization of an object in space with respect to one's body depends on knowledge of body, head, and eyeball position. If the head and body are fixed, the only variable is position of the eyeball.

There is evidence that the position of the eyeball is *not* adequately known on the basis of proprioceptive signals from the extraocular muscles. Hence, with head and body in a fixed position, accuracy of localizing an object in space would be poor if the only information about eyeball position came from such proprioceptive signals.

There is evidence that smooth tracking movements of the eye are controlled and directed by the direction and velocity of movement across the retina and *not* by target location. Saccadic eye movements, on the other hand, are directed on the basis of target location on the retina. Hence, if a target is fixated by means of a saccadic movement, efferent signals relevant to target location would have been issued. If a target is tracked by a smooth eye movement, however, efferent signals concerning direction and velocity of movement would have been issued—information not optimally useful for knowing the target location.

Consequently, if a record of efferent signals is available, localization in space of a target should be better following fixation by a saccadic eye movement than following a smooth tracking eye movement. Having found this result, we regard it as evidence for the existence of information based on this hypothesized record of efferent signals.

It is, of course, possible that there are alternative interpretations of the data we have presented. No such plausible alternatives occur to us, how-

ever. It does not, for example, seem possible to maintain any alternative interpretations in terms of confusion introduced by the tracking procedure, since it is clear, in the "head-movement condition," that the tracking procedure, in and of itself, does not interfere with accuracy.

Another possible alternative explanation could be elaborated as follows. Presumably, during the period of darkness between trials, the subject's eyes revert to some "normal" frontal position. Such a normal position is probably a reference point for location in the visual field, and, presumably, directions are issued to the extraocular muscles with respect to some such reference point. The eyeball then moves, in accordance with the efferent directions, in a saccadic, ballistic movement. Under such circumstances the initial movement of the eye to fixate the target is *not* a continuously controlled movement. Once started it proceeds to its destination. The saccadic movement, hence, must have a complete set of directions issued at the beginning.

It thus becomes clear that, in order to issue directions that are relatively accurate for the initial ballistic movement of the eye, information as to the location in space of the target must exist before the directions are issued. And indeed, this information must be obtained on the basis of the stimulation of the periphery of the retina when, with the eyes in frontal position, the target light suddenly appears. It is on the basis of this information that the initial ballistic eye movement is more or less accurately directed.

Why, then, is it necessary to say that the differences obtained between the "eye-movement-only" conditions are due to a record of the efferent impulses *actually* sent out to the muscles? Why could we not simply maintain

that the information the person has as to the location of the target light is simply the information on the basis of which the efferent directions were issued? After all, on the tracking trials the subject did not see the target light at its final position in peripheral vision. The results of the "head-movement" conditions rule out this explanation of the results. If seeing the target light in peripheral vision were important, the tracking conditions should still be inferior even with head movements allowed.

One must admit, however, that information based on a record of the efferent signals is not likely to be better than the information on the basis of which those efferent signals were sent. Our present data cannot answer questions concerning the relation between these two things. Our experiment does, however, confirm the existence, and usefulness, of outflow information.

REFERENCES

BRINDLEY, G. S., & MERTON, P. A. The absence of position sense in the human eye. *Journal of Physiology*, 1960, 153, 127-130.

- FENDER, D. H. The eye-movement control system: Evolution of a model. In R. F. Reiss (Ed.), *Neural theory and modeling*. Stanford: Stanford Univer. Press, 1964. Pp. 306-324.
- GRAEFE, A. VON. *Handbuch der gesamten Augenheilkunde*, 1878, 6, 18-21.
- HELMHOLTZ, H. VON. *Treatise on physiological optics*. (3rd ed.) (Ed. & trans. by P. C. Southall) Vol. 3. Menasha, Wis.: Optical Society of America, 1925.
- HOLST, E. VON. Relations between the central nervous system and the peripheral organs. *British Journal of Animal Behavior*, 1954, 2, 89-94.
- JAMES, W. *Principles of psychology*. Vol. 2. New York: Dover, 1950.
- MACH, E. *The analysis of sensations*. Chicago: Open Court, 1914.
- MERTON, P. A. Absence of conscious position sense in the human eyes. In M. B. Bender (Ed.), *The oculomotor system*. New York: Harper & Row, 1964. Pp. 314-320.
- RASHBASS, C. The relationship between saccadic and smooth tracking eye movements. *Journal of Physiology*, 1961, 159, 326-338.
- WHITTERIDGE, D. Afferent mechanisms in the initiation and control of eye movement. In *Proceedings of the International Union of Physiological Science: XXII International Congress*. Amsterdam: Excerpta Medica Foundation, 1962. Pp. 509-512.

(Received June 10, 1964)

CENTRECEPHALIC THEORY AND INTERHEMISPHERIC TRANSFER OF VISUAL HABITS¹

ROBERT THOMPSON

Louisiana State University

Recent anatomical and psychological data suggest that visual pattern-discrimination habits are mediated by a direct occipito-mesencephalic tract, while a simple brightness-discrimination habit is mediated by an occipito-preecto-tegmental pathway. Both of these projections are homolateral. By assuming that the memory trace develops at the terminal endings of the occipito-fugal pathway subserving the learned response, it is possible to explain the presence or absence of inter-ocular transfer in split-brain animals. This scheme represents an extension of Penfield's centrencephalic theory and a defense for the possible existence of a subcortically induced memory trace.

In recent years, research on the "split-brain" animal has revealed an important function of the corpus callosum in visual discrimination learning (Myers, 1961; Sperry, 1961). This function has to do with intercommunication between the two hemispheres of learning experiences restricted to only one side of the brain. More generally, the data emerging from split-brain animals have been interpreted (Bureš & Burešová, 1960; Russell & Ochs, 1963) as inimical to a centrencephalic theory (Penfield, 1954a) and to any other theory having as one of its tenets the subcortical origin of the memory trace (e.g., Gastaut, 1958). Myers and Sperry (1958) expressed a similar view:

These facts point to cortex (and associated thalamic nuclei) as the probable site of elaboration of visual sensory input. The contribution of brain-stem mechanisms to such high-level, "psychic" activity has yet to be fully explored. Present evidence, however, points toward involvement of brain stem in preparing and maintaining levels of activity in the cortex necessary for such

high-level functions, rather than toward a primary involvement in these functions. Accordingly, it is assumed in the following discussion that the memory mechanisms underlying the involved pattern discrimination responses reside somewhere within the cortical gray mantle, or at least can be grossly disorganized by appropriate cortical insult [p. 301-302].

In the current paper, some recent anatomical and psychological data will be discussed in relation to Penfield's centrencephalic theory. Taken as a whole, these data strengthen Penfield's position that occipito-fugal pathways terminating within the brain stem are functionally significant in visually guided behavior. By assuming that the memory trace develops at the terminal endings of these cortico-fugal projections, it is possible to incorporate the findings of interhemispheric transfer of visual learning within the framework of a centrencephalic model of brain function. Contrary to current views, the absence of transfer of a pattern discrimination in split-brain animals will be shown to argue more strongly for a subcortically induced memory trace than it does for a cortically induced memory trace.

¹ Aided by Grant MH-08377-01 from the National Institute of Mental Health, United States Public Health Service.

PENFIELD'S CENTRECEPHALIC THEORY

One of the principal proponents of the view that subcortical processes take precedence over cortical processes in the integration of the total activity of the brain is Penfield (1952, 1954a, 1954c, 1958). According to Penfield, there exists within the brain stem (portions of the diencephalon, mesencephalon, and probably the metencephalon) an ensemble of nuclei and circuits which serves to coordinate and integrate the activities of the cerebral hemispheres. This centrally located, integrating ensemble is termed the "centrencephalic system" (Penfield, 1952). Functionally, it represents

a ganglionic area in which that stream of nervous impulses must arise that produces voluntary activity, an area in which the sensory pathways culminate in neurone circuits and in which the information relative to past experience is made available, an area in which those nervous mechanisms are to be found which are prerequisite to the existence of intellectual activity and prerequisite to the initiation of the patterned stream of efferent impulses that produce the planned action of the conscious man [Penfield, 1954c, p. 286].

Thus, with respect to visually guided behavior, Penfield conceives visual impulses as being relayed to the centrencephalic system once they reach the occipital cortex. By virtue of this system's vast interconnecting circuitry, engrams relative to this visual input are excited and compared with present experience. On the basis of this information, the system discharges a particular pattern of "volitional" impulses to cortical and subcortical motor mechanisms.

The evidence that Penfield cites in support of the subcortical origin of the central integrating system is largely composed of clinical observations. Particularly relevant to Penfield's position are the well-established findings

that small interferences with the brain stem either by undue pressure, injury, or local epileptic discharge abolish consciousness and purposeful action, while extensive cortical excisions leave these "psychical" activities intact (Penfield, 1957). Another source of evidence focuses on the manner in which the precentral motor gyrus is activated. According to Penfield (1954a), it is untenable to envisage the motor area as being functionally excited by transcortical pathways since ablation of cortical tissue surrounding the precentral gyrus fails to interfere with the production of skilled movements. The only alternative is that the motor cortex is under control of a ganglionic mass residing beneath the cortical gray mantle. Consistent with this possibility are the findings relative to the local abolishment of precentral beta rhythms during voluntary action (Penfield & Jasper, 1954).

In his earlier writings, Penfield (1954b) concluded that the temporal cortex was the site of the memory trace. This conclusion has since been withdrawn in view of the absence of major amnesic effects following bilateral temporal resections in humans (Penfield, 1959). While Penfield does not explicitly advocate a subcortical memory trace, he does give some hint of this belief in his "Horowitz" lecture. In discussing speech mechanisms, Penfield (1963) stated that:

This voluntary outflow is not transcortical from speech cortex to motor cortex as was sometimes assumed. It is through the centrencephalic system. And the "hook-up" between non-verbal concept and word-idea must have been through that system [p. 45].

SOME RECENT SUPPORTIVE DATA

It will be noted that Penfield has not specifically addressed his theory to the field of animal learning. It was originally adopted as a working hypothesis

with specific application to clinical neurological cases. One must hasten to add, however, that centrencephalic theory stands in good stead with recent neuroanatomical findings (Russell, 1961) and receives considerable support from animal studies relative to visuomotor behavior (Myers, Sperry, & McCurdy, 1962).

Research from our laboratory on the albino rat not only points to the existence of a centrencephalic mechanism underlying visual discrimination learning, but sheds some light on possible functional pathways linking the occipital cortex with the brain stem (Thompson, Rich, & Langer, 1964). During the past 8 years, my associates and I have been surgically removing portions of the brain in previously trained rats in an effort to locate those nuclei and pathways which maintain a brightness-discrimination habit. During the course of this research, virtually every part of the tel-, di-, mes-, and metencephalon has been examined as to its contribution to the performance of the discriminative response. The only lesions which consistently yielded positive effects include the visual cortex, pretectum (anterior region of the *nucleus posterior thalami* of Gurdjian, 1927), lateral posterior hypothalamus, and ventral mesencephalon (Thompson, 1963a). That these data suggest a centrencephalic mechanism is indicated by the very nature of the anatomical connections existing between these critical structures. First of all, there is general agreement that the occipital cortex of rodents projects, in part, to the nucleus posterior (Combs, 1949, 1951; d'Hollander, 1922). Secondly, the nucleus posterior has been shown to send a fiber projection to the ventral mesencephalon (Krieg, 1947; Papez & Freeman, 1930). Finally, the lateral-hypothalamic and ventral-mesencephalic areas

are anatomically interrelated (Crosby & Woodburne, 1951; Krieg, 1932) and very probably contain the ganglionic groups and circuits necessary for the integration and execution of locomotor behavior (Hinsey, Ranson, & McNattin, 1930; Keller, 1932; Waller, 1940), the defence reaction (Abrahams, Hilton, & Zbrozyna, 1960; Bard, 1928), and other somatic behavior patterns (Woods, 1964). Taken as a whole, these data suggest that the performance of a discriminative response to stimuli differing in brightness requires a relay of the visual input to the brain stem by an occipito-pretecto-mesencephalic pathway.²

CORTICOFUGAL PATHWAYS MEDIATING PATTERN- AND BRIGHTNESS- DISCRIMINATION HABITS

According to Penfield's theory, visual impulses are channeled to certain regions of the brain stem once they have reached the occipital cortex. While several recognizable pathways are available to effect this relay (see Meikle & Sprague, 1964), only two seem to be functionally significant in discrimination learning. One of these pathways has already been mentioned. It is indirect and extends from the occipital cortex to the ventral mesencephalon by way of the pretectal area. The first component of this pathway, the occipito-pretectal projection, has been identified in all mammals studied (Altman, 1962; d'Hollander, 1922; Nauta & Bucher, 1954). The second component, the so-called "pretectotegmental tract," is found in submammalian species (Clark, 1932) as well

² In a recent paper, Rich and Thompson (1965) outlined a centrencephalic mechanism underlying the performance of a conditioned avoidance response. Unlike the conventional visual discrimination habit, the avoidance habit is highly dependent upon the limbic system (Thompson et al., 1964).

as in the opossum (Bodian, 1940; Tsai, 1925), the armadillo (Papez, 1932), the rat (Krieg, 1947; Papez & Freeman, 1930), the cat (Bucher & Bürgi, 1952), the monkey (Clark, 1932), and even the human (Kuhlenbeck & Miller, 1949). The majority of these reports disclose that the pretectotegmental tract terminates in the region of the red nucleus and *substantia nigra*.

The second corticofugal pathway which functionally relates the occipital cortex with the brain stem is a direct one. This occipito-mesencephalic tract has recently been demonstrated in rats (Valverde, 1962) and cats (Pearce, 1960). It very probably exists in monkeys as well (Jasper, Ajmone-Marsan, & Stoll, 1952; Mettler, 1935). In the rat, the projection is chiefly to the basal half of the midbrain tegmentum, while the projection in cats seems to be limited to the dorsal half (the subcollicular area). This corticofugal projection is homolateral, a finding to be elaborated upon later in the discussion of the behavior of split-brain animals.

Several lines of evidence suggest that the direct occipito-mesencephalic tract functions primarily in pattern-discrimination learning, while the occipito-pretectotegmental pathway functions chiefly in simple brightness-discrimination learning. First of all, lesions of the pretectum do not prevent rats from relearning a visual pattern-discrimination habit (Thompson & Rich, 1963). Ablation of the occipital cortex, on the other hand, virtually eliminates the capability of rats (Lashley & Frank, 1934; Schwartz & Clark, 1957), cats (Myer, 1963; Smith, 1938), and dogs (Marquis, 1934) to perform successfully on a pattern discrimination. Secondly, pretectal damage has a significantly greater disruptive effect on a simple brightness discrimination than on a pattern discrimination (Thompson

& Rich, 1963). Finally, cats subjected to lesions in the subcollicular region of the midbrain to which the occipito-mesencephalic tract terminates have considerable difficulty in the performance of visual pattern discriminations (Blake, 1959; Myers, 1964; Sprague, Chambers, & Stellar, 1961) but successfully retain a simple brightness discrimination (Blake, 1959).

Further evidence in support of this functional dichotomy is afforded by the following experiment recently performed in our Louisiana laboratory. Adult albino rats were trained preoperatively on either a simple brightness (black card *versus* white card) discrimination or a pattern (vertically striped card *versus* horizontally striped card) discrimination. Since the apparatus, training, surgical, and histological procedures were quite comparable to those reported previously (Thompson & Rich, 1963), it is possible to dispense with a detailed description of the investigation. Following learning, the majority of animals were subjected to bilateral electrolytic lesions of the caudal mesencephalon. In some cases, a special effort was made to damage the basolateral midbrain area (*nucleus cuneiformis*) to which the direct occipito-mesencephalic tract terminates (Valverde, 1962). After a recovery period ranging from 2 to 3 weeks, the control and operated rats were required to relearn the discrimination habit that was learned preoperatively. Individual retention scores were expressed in terms of percentage of error savings, and the significance of group differences calculated by a non-parametric test (Festinger, 1946).

Table 1 summarizes the original learning and retention scores for the various groups differentiated on the basis of the site of the lesion. A total of 16 rats initially trained on the pattern discrimination received bilateral

TABLE 1
MEAN LEARNING AND RETENTION SCORES FOR ALL GROUPS

Group	Pattern discrimination			Brightness discrimination		
	N	Learning Errors	Retention savings (%)	N	Learning Errors	Retention savings (%)
Control	6	23.8	97.5	4	7.5	100
Basolateral tegmentum	4	17.8	-37.5 ^a	5	8.8	89.0
Lateral tegmentum	6	20.2	93.3			
Subcollicular area	3	14.3	100			
Intercollicular area	3	18.3	100			

^a Significantly different from the controls at the .01 level.

midbrain lesions. The ventral portion of the *nucleus cuneiformis* along with a caudolateral segment of the *substantia nigra* were damaged in four animals. Two of these animals were unable to relearn the task even after having received twice the number of trials required preoperatively. The remaining two animals earned savings scores of -10% and 50%. As a group, these four operated rats were inferior to the six controls at the .01 level of significance. Figure 1 (Parts A, B, and C) illustrates the lesion of one rat which failed to relearn the pattern discrimination. Those two animals which succeeded in reacquiring the habit suffered appreciably smaller lesions of the *nucleus cuneiformis* than those which failed to relearn. It is interesting to note that at the time of the postoperative test, these four animals appeared quite normal in every respect. No obvious motor or sensory deficits were ever observed. The only disturbance appearing in these four animals was an aphagia which lasted for the first 3 postoperative days. An additional six animals sustained lesions of the lateral tegmentum which damaged, in part, the dorsal portion of the *nucleus cuneiformis*. All earned savings scores in excess of 83%. Finally, six rats subjected to extensive

destruction of the subcollicular area or the intercollicular zone earned perfect retention scores. Figure 1 (Parts D, E, and F) summarizes the various sites of the lesions examined in this experiment. With respect to these three levels of the midbrain, only damage to the basolateral tegmentum yields significant effects on retention.

That the ventral portion of the *nucleus cuneiformis* appears to function specifically for pattern-discrimination tasks is indicated by the performance of five rats trained preoperatively on the simple brightness discrimination. These five animals sustained lesions similar to those shown in the left column of Figure 1. Three achieved perfect savings scores, while the remaining two earned scores of 67% and 78%. As a group, these five operated animals were not significantly inferior to the control group. It is noteworthy to report that the two animals which earned subperfect retention scores sustained midbrain lesions which were considerably larger than those suffered by the two rats which failed to relearn the pattern discrimination.

Before leaving this section, three comments must be made in connection with these corticofugal pathways involved in visual discrimination learning. The first comment is concerned

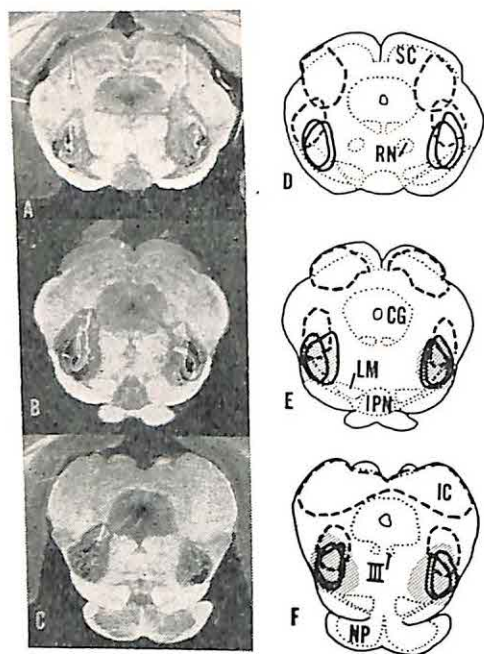


FIG. 1. Photographs and camera drawings of frontal sections showing lesion placements. (Parts A, B, and C represent photographs of unstained sections showing a basolateral midbrain lesion which prevented relearning of a pattern discrimination in Rat 11M. Parts D, E, and F represent drawings of corresponding midbrain levels showing composites of lesions investigated in this experiment. Stippled areas, Parts E and F, indicate the rostral midbrain projection of the occipito-mesencephalic tract. Areas enclosed by heavy solid lines denote positions of lesions preventing relearning. Areas enclosed by heavy interrupted lines denote positions of lesions yielding no loss in retention. The enclosed lesioned areas represent regions of total destruction only. Abbreviations: CG, central gray; IC, inferior colliculus; IPN, interpeduncular nucleus; LM, medial lemniscus; NP, pons; RN, red nucleus; SC, superior colliculus; III, nucleus of third nerve.)

with the contribution of the direct occipito-fugal pathway in learning a difficult brightness discrimination. It now appears that the performance of a difficult (near-threshold) brightness habit may involve the utilization of the direct occipito-mesencephalic tract to a greater extent than the indirect path-

way. This conclusion is based upon the findings that rats (Bauer & Cooper, 1964; Lashley, 1930) and cats (Smith, 1937) lacking the striate cortex exhibit increasing deficits in performance when the difference between the brightness discriminanda is diminished. Rats with pretectal lesions, in contrast, exhibit no greater deterioration on near-threshold brightness discriminations as compared with suprathreshold brightness discriminations (Thompson & Rich, 1963). In fact, pretectal lesions seem to produce more marked effects on a black-white discrimination (Thompson & Massopust, 1960) than on a medium gray-white discrimination (Thompson & Rich, 1963).

The second comment has to do with the dispensability of the occipito-pretectal projection in the performance of certain simple visual habits. It is well established that animals without the striate cortex can learn a simple brightness discrimination as fast as normals (Lashley, 1929; Marquis, 1934; Smith, 1937). Similarly, total ablation of the occipital cortex may not seriously affect the performance of either a discrimination habit based upon differences in luminous flux (Bauer & Cooper, 1964) or an avoidance-conditioned response elicited by a visual signal (Marquis & Hilgard, 1936; Meyer, 1963; Thompson, 1960; Wing & Smith, 1942). Since massive pretectal damage may destroy the rat's ability to develop simple brightness discriminations (Layman, 1936; Thompson & Massopust, 1960) and avoidance responses to visual stimuli (Thompson, 1963b), the pretecto- tegmental tract may, under certain conditions, mediate visual habits independent of the corticofugal projection. This possibility is strengthened by the observation that striate rats, upon learning a simple brightness discrimination, exhibit a significant retention deficit

following bilateral damage to the pretectum (Thompson & Rich, 1963). In this case, the visual input to the pretectum may be transmitted either by direct optic fibers terminating within the pretectal area or by indirect association fibers originating in the lateral geniculate or tectal nuclei (Meikle & Sprague, 1964).

The final comment deals with the possible mechanisms underlying the functional dichotomy of the occipito-fugal systems. According to the recent work of Hubel and Wiesel (1962), most of the cells composing the striate cortex are maximally responsive to patterns of light stimuli (lines and forms) and minimally responsive to changes in illumination. It is only a short extrapolation to the assumption that those cells responsive to patterned stimuli excite almost exclusively the direct occipito-mesencephalic tract. Thus, this tract may mediate form-discrimination habits by virtue of its inherent connections with particular occipital cells. Since changes in general illumination are relatively ineffective in evoking cortical responses, these stimuli are assumed to activate primarily the pretecto-tegmental tract. The afferent limb of this pathway may very probably be the same as that involved in the light reflex (Ranson & Magoun, 1933). This pretecto-tegmental tract, therefore, will mediate conditioned responses elicited by changes in the general level of illumination. With respect to visual intensity discrimination, no clear-cut electrophysiological data are available to suggest possible correlates. However, it is not inconceivable that the organization of the visual cortex is such that the occipito-pretecto-tegmental pathway is activated by striate cells sensitive to gross flux differences, and the occipito-mesencephalic tract is excited by striate cells sensitive to slight flux differences. As a consequence,

suprathreshold and near-threshold brightness-discrimination habits would be subserved by different occipito-fugal pathways.

Let us now turn to some data which are customarily considered to be at variance with a centrecephalic theory.

INTERHEMISPHERIC TRANSFER EXPERIMENTS

A normal cat trained monocularly on a visual pattern-discrimination problem will exhibit excellent transfer of learning when the untrained eye is utilized and the trained eye is covered. This transfer is also displayed in cats in which the crossed optic fibers have been cut (Myers, 1955). If, however, the corpus callosum and the crossed optic fibers are severed, no transfer is observed (Myers, 1961; Sperry, Stamm, & Miner, 1956).

These data seemingly detract from Penfield's theory in two respects. First, Penfield assumes that the activities of the cerebral hemispheres are coordinated and integrated through symmetrical ascending connections of the brain stem. This coordination and integration is conceived to be independent of the corpus callosum. Numerous experiments have recently been reported which attest to the coordinating function of the callosal system (Myers, 1961; Sperry, 1961). Secondly, if one assumes correctly that Penfield subscribes to a subcortical memory trace, then Penfield's centrecephalic theory demands that the memory trace should be bilaterally represented within the brain stem. This follows from the assumption that each hemisphere has *bilateral descending connections with the brain stem*. However, the absence of interocular transfer of pattern-discrimination learning in split-brain animals makes it clear that a bilateral subcortical engram is not formed.

Concerning the first point, Penfield must accept the fact that callosal fibers do function under certain conditions to interrelate the activities of the cerebral hemispheres. But this fact does not in any way speak against further coordinating and integrating functions of the brain stem. It has been shown, for example, that callosal fibers do not function critically in connection with interocular transfer of a simple brightness-discrimination habit (Meikle, 1960; Meikle & Sechzer, 1960). Additional studies demonstrate that the callosal system does not mediate learned responses in which the visual input is confined to one hemisphere, while the limb performing the responses is governed by the contralateral hemisphere (Voneida, 1963). It has also been reported that the split-brain monkey can perform a visuo-tactile task in which the tactile stimulus is restricted to one hemisphere while the visual stimulus is restricted to the opposite hemisphere (Sperry, 1961). Furthermore, Doty (1961) has summarized several experiments which show that the callosal pathway is unnecessary in the development of conditioned responses when the CS and UCS are applied to opposite hemispheres. The integrating circuits carrying out the foregoing activities must be subcortical in origin.³ Thus, interhemispheric integration through the corpus callosum weakens a centrencephalic position only to the ex-

tent that the *brain stem does not provide all of the necessary circuitry* to coordinate the activities of the two hemispheres.

Concerning the second point, while the work of Myers, Sperry, and others strongly suggests that the memory trace for visual pattern discrimination develops unilaterally in the split-brain animal, considerably more data are needed to establish that the memory trace for visual habits is formed within the cerebral cortex rather than within the brain stem. To cite some negative evidence, the memory trace for an avoidance-conditioned response to a visual signal does not reside within the occipital cortex (Marquis & Hilgard, 1936; Thompson, 1960; Wing & Smith, 1942). Schwartz and Clark (1957) have found that the memory trace for a discrimination between a flickering light and a steady light is not formed within the visual cortex. Recently, Bauer and Cooper (1964) reported that a discriminative response based upon luminous flux survives complete bilateral removal of the posterior half of the cerebrum. Finally, cats (Hernández-Peón & Brust-Carmona, 1961; Meyer, 1963) and rats (Thompson, 1959) deprived of most of their neocortex learn certain tasks as fast as normals. The claim that the engram for simple habits is localized at subcortical levels while the cortical gray mantle retains the engram for complex habits is also in need of experimental verification.

In the succeeding account, an interpretation of the split-brain data will be presented within the framework of a centrencephalic theory. This interpretation is based upon the information concerning the functional characteristics of the two occipito-fugal pathways discussed previously and upon an assumption dealing with the specific subcortical site of the engram.

³ Since this paper was completed, Sechzer (1964) reported that split-brain cats show significant interocular transfer of a visual pattern discrimination under the condition of shock-avoidance motivation. This finding suggests that the occipito-pretecto-tegmental pathway is potentially capable of mediating visual pattern discriminations. Interestingly, Krieg (1947) has reported that the pretecto-tegmental tract can be traced to the level of the pons wherein lies the *nucleus cuneiformis*.

AN INTERPRETATION

It will be recalled that the direct occipito-mesencephalic tract is homolateral (Pearce, 1960; Valverde, 1962). That is to say, the left occipital cortex projects to the left mesencephalon, and vice versa. At this point, if it is assumed that *the engram forms at the terminal endings of the corticofugal pathway subserving the learned habit*, then we can explain in terms of a centrencephalic theory the presence or absence of transfer of certain visual habits in the split-brain animal. Consider, first of all, the split-brain cat that is initially trained on a pattern discrimination with the left eye only. Performance of this habit would be mediated by the left occipito-mesencephalic tract, and the memory trace would be formed within the left mesencephalon. When the right eye is exposed to the problem and the left eye covered, only the right occipito-mesencephalic tract is activated. No transfer of learning would be predicted. The only way for transfer to take place is by channeling the visual input to the left occipito-mesencephalic tract. This can only be effected if the crossed optic fibers or if the callosal fibers are intact. Figure 2 (Part A) illustrates this phenomenon.

It is readily apparent that this scheme provides for the possibility that split-brain animals can acquire antagonistic visual habits with the two eyes (Myers, 1956). The left occipito-mesencephalic tract could be utilized to mediate one habit with the left eye, while the right occipito-mesencephalic tract could be utilized to mediate an altogether different habit with the right eye.

It has been reported that near-threshold brightness discriminations will not transfer in split-brain animals (Meikle, 1960). As pointed out ear-

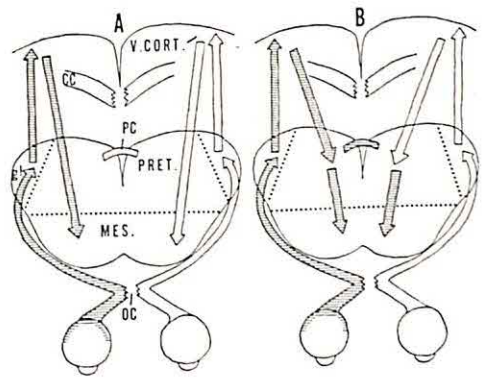


FIG. 2. A semidiagrammatic representation of the split-brain animal showing the critical neuronal pathways conveying optic impulses to the brain stem. (The shaded pathways indicate the route over which visual impulses travel when the left eye is exposed. See text for a further description. Part A: Neuronal pathways subserving pattern- and near-threshold brightness-discrimination habits. Part B: Neuronal pathways subserving a simple brightness-discrimination habit. Abbreviations: CC, corpus callosum; gl, lateral geniculate nucleus; MES., mesencephalon; OC, optic chiasma; PC, posterior commissure; PRET., pretectal area; V. CORT., occipital cortex.)

lier, a near-threshold brightness-discrimination habit is very probably mediated by the same direct occipito-fugal pathway that is utilized in pattern-discrimination habits. Thus, no transfer would be predicted.

Let us now consider the simple brightness-discrimination habit which does transfer in the split-brain animal (Meikle, 1960; Meikle & Sechzer, 1960). It will be remembered that a simple brightness discrimination is probably mediated by the occipito-pretecto-tegmental pathway. The engram, in this case, would develop at the terminal endings of the pretecto-tegmental tract (the region of the red nucleus and *substantia nigra*). While this projection is homolateral (Krieg, 1947; Papez & Freeman, 1930), the left pretectum communicates with the right pretectum through the posterior

and tectal commissures (Bucher & Bürgi, 1952; Gurdjian, 1927; Kuhlenbeck & Miller, 1942). Thus, a split-brain animal trained on a supra-threshold brightness-discrimination habit with the left eye will utilize the left occipito-preecto-tegmental pathway and, by way of the subcortical commissures, the right preecto-tegmental tract. When the right eye is now exposed to the problem, the right occipito-preectal tract is activated which, in turn, excites both the left and right preecto-tegmental tracts. Transfer of simple brightness discriminations would therefore be predicted. Figure 2 (Part B) illustrates this phenomenon.

It would follow that sectioning of the posterior and tectal commissures as well as the crossed optic and callosal fibers should prevent transfer of simple brightness habits. Partial confirmation of this prediction comes from the work of Voneida (1963) who failed to obtain transfer of an avoidance response to light only when the posterior and tectal commissures were cut in addition to the optic and callosal fibers.

OTHER CONSIDERATIONS

The basic assumptions of the foregoing account are not restricted to visual discrimination learning. In the cat, Pearce (1960) has found that the sensory-motor cortex has a direct homolateral projection to the ventral mesencephalon. A similar direct projection has recently been observed in the albino rat (Valverde, 1962). By making the claim that this pathway mediates somesthetic discriminations and that the engram develops at the terminal endings of this corticofugal projection, then it is possible to explain the absence of transfer of somesthetic discrimination habits in cal-

losum-sectioned animals (Ebner & Myers, 1962; Stamm & Sperry, 1957). In a similar fashion, it is possible to explain the lateralization of the memory trace in animals required to learn a simple conditioned response during unilateral spreading depression of one hemisphere (Bureš & Burešová, 1960).

It is noteworthy to mention the ease with which these assumptions can be tested. In some respects, the nervous structures and circuits proposed to underlie visually guided behavior have been defined even more precisely than originally intended. The luster of this precision, however, is marred to the extent that the temporal cortex has not been fitted into this anatomical scheme. The specific role played by the temporal lobes in visual discrimination performance must await further elucidation of the functional pathways proceeding to and from this section of the cerebral cortex. The research plan followed by Chow (1961) may be fruitful in this regard. In any event, the functional significance of the temporal cortex would be predicted to emerge from anatomical relationships either with the occipital cortex, preectum, or brain stem.

Finally, a remark must be made in connection with possible phylogenetic differences in functional pathways mediating discriminative responses. Throughout this paper, it has been assumed that the anatomical mechanisms underlying the performance of discrimination tasks are *essentially the same in all mammalian forms*. This assumption is basic to a centrencephalic theory which views the phylogenetic new brain (neocortex) as being an extension of the phylogenetic old brain (brain stem). Despite the vast differences in anatomical complexity, the functional interrelationships between the telencephalon and the dimesen-

cephalic zone are conceived to be preserved throughout the mammalian series. The anatomical review of the corticofugal pathways is not inconsistent with this assumption. Furthermore, it may be more than coincidence that an investigation of the brain-behavior relationships in the rodent has led to the same conclusions as those drawn by Penfield's analysis of the neural mechanisms underlying human behavior.

REFERENCES

- ABRAHAMS, V. C., HILTON, S. M., & ZBROZYNA, A. Active muscle vasodilatation produced by stimulation of the brain stem: Its significance in the defence reaction. *Journal of Physiology*, 1960, 154, 491-513.
- ALTMAN, J. Some fiber projections to the superior colliculus in the cat. *Journal of Comparative Neurology*, 1962, 119, 77-96.
- BARD, P. A diencephalic mechanism for the expression of rage with special reference to the sympathetic nervous system. *American Journal of Physiology*, 1928, 84, 490-515.
- BAUER, J. H., & COOPER, R. M. Effects of posterior cortical lesions on performance of a brightness-discrimination task. *Journal of Comparative and Physiological Psychology*, 1964, 58, 84-92.
- BLAKE, L. The effect of lesions of the superior colliculus on brightness and pattern discrimination in the cat. *Journal of Comparative and Physiological Psychology*, 1959, 52, 272-278.
- BODIAN, D. Studies on the diencephalon of the Virginia opossum: II. The fiber connections in normal and experimental material. *Journal of Comparative Neurology*, 1940, 72, 207-298.
- BUCHER, V. M., & BÜRG, S. M. Some observations on the fiber connections of the di- and mesencephalon in the cat: II. Fiber connections of the pretectal region and the posterior commissure. *Journal of Comparative Neurology*, 1952, 96, 139-177.
- BUREŠ, J., & BUREŠOVÁ, OLGA. The use of Leão's spreading depression in the study of interhemispheric transfer of memory traces. *Journal of Comparative and Physiological Psychology*, 1960, 53, 558-563.
- CHOW, K. L. Anatomical and electrophysiological analysis of temporal neocortex in relation to visual discrimination learning in monkeys. In J. F. Delafresnaye (Ed.), *Brain mechanism and learning*. Oxford: Blackwell, 1961. Pp. 507-523.
- CLARK, W. E. The structure and connections of the thalamus. *Brain*, 1932, 55, 406-470.
- COMBS, C. M. Fiber and cell degeneration in the albino rat after hemidecortication. *Journal of Comparative Neurology*, 1949, 90, 373-401.
- COMBS, C. M. The distribution and temporal course of fiber degeneration after experimental lesions in the rat brain. *Journal of Comparative Neurology*, 1951, 94, 123-175.
- CROSBY, E. C., & WOODBURN, R. T. The mammalian midbrain and isthmus regions: Part II. The fiber connections: C. The hypothalamo-tegmental pathways. *Journal of Comparative Neurology*, 1951, 94, 1-32.
- D'HOLLANDER, F. Recherches anatomiques sur les couches optiques. *Archives de Biologie*, 1922, 32, 249-344.
- DOTY, R. W. The role of subcortical structures in conditioned reflexes. *Annals of New York Academy of Sciences*, 1961, 92, 939-945.
- EBNER, F. F., & MYERS, R. E. Corpus callosum and the interhemispheric transmission of tactile learning. *Journal of Neurophysiology*, 1962, 25, 380-391.
- FESTINGER, L. The significance of difference between means without reference to frequency distribution function. *Psychometrika*, 1946, 11, 97-105.
- GASTAUT, H. The role of the reticular formation in establishing conditioned reactions. In H. H. Jasper et al. (Eds.), *Reticular formation of the brain*. Boston: Little, Brown, 1958. Pp. 561-579.
- GURDJIAN, E. S. The diencephalon of the albino rat: Studies on the brain of the rat. *Journal of Comparative Neurology*, 1927, 43, 1-114.
- HERNÁNDEZ-PEÓN, R., & BRUST-CARMONA, H. Functional role of subcortical structures in habituation and conditioning. In J. F. Delafresnaye (Ed.), *Brain mechanisms and learning*. Oxford: Blackwell, 1961. Pp. 393-408.
- HINSEY, J. C., RANSON, S. W., & McNATTIN, R. F. The role of the hypothalamus and mesencephalon in locomotion. *Archives of Neurology and Psychiatry*, 1930, 23, 1-43.
- HUBEL, D. H., & WIESEL, T. N. Receptive fields, binocular interaction and functional

- architecture in the cat's visual cortex. *Journal of Physiology*, 1962, **160**, 106-154.
- JASPER, H., AJMONE-MARSAN, C., & STOLL, J. Corticofugal projections to the brain stem. *Archives of Neurology and Psychiatry*, 1952, **67**, 155-166.
- KELLER, A. D. Autonomic discharges elicited by physiological stimuli in mid-brain preparations. *American Journal of Physiology*, 1932, **100**, 576-586.
- KRIEG, W. J. S. The hypothalamus of the albino rat. *Journal of Comparative Neurology*, 1932, **55**, 19-89.
- KRIEG, W. J. S. Connections of the cerebral cortex: I. Albino rat: C. Extrinsic connections. *Journal of Comparative Neurology*, 1947, **86**, 267-394.
- KUHLENBECK, J., & MILLER, R. N. The pretectal region of the rabbit's brain. *Journal of Comparative Neurology*, 1942, **76**, 323-365.
- KUHLENBECK, J., & MILLER, R. N. The pretectal region of the human brain. *Journal of Comparative Neurology*, 1949, **91**, 369-407.
- LASHLEY, K. S. *Brain mechanisms and intelligence*. Chicago: Univer. Chicago, 1929.
- LASHLEY, K. S. The mechanism of vision: I. The influence of cerebral lesions upon the threshold of discrimination for brightness in the rat. *Journal of Genetic Psychology*, 1930, **37**, 461-480.
- LASHLEY, K. S., & FRANK, M. The mechanism of vision: X. Postoperative disturbances of habits based on detail vision of the rat after lesions in the cerebral visual areas. *Journal of Comparative Psychology*, 1934, **17**, 355-391.
- LAYMAN, J. D. Functions of the superior colliculi in vision. *Journal of Genetic Psychology*, 1936, **49**, 33-47.
- MARQUIS, D. G. Effects of removal of the visual cortex in mammals, with observations on the retention of light discrimination in dogs. *Research Publications in the Association of Nervous and Mental Disorders*, 1934, **13**, 558-592.
- MARQUIS, D. G., & HILGARD, E. R. Conditioned lid responses to light in dogs after removal of the visual cortex. *Journal of Comparative Psychology*, 1936, **22**, 157-178.
- MEIKLE, T. H. Role of corpus callosum in transfer of visual discriminations in the cat. *Science*, 1960, **132**, 1496. (Abstract)
- MEIKLE, T. H., & SECHZER, J. A. Interocular transfer of brightness discrimination in "split-brain" cats. *Science*, 1960, **132**, 734-735.
- MEIKLE, T. H., & SPRAGUE, J. M. The neural organization of the visual pathways in the cat. *International Review of Neurobiology*, 1964, **6**, 149-189.
- METTLER, F. A. Corticofugal fiber connections of the cortex of *Macaca mulatta*: The occipital region. *Journal of Comparative Neurology*, 1935, **61**, 221-256.
- MEYER, P. M. Analysis of visual behavior in cats with extensive neocortical ablations. *Journal of Comparative and Physiological Psychology*, 1963, **56**, 397-401.
- MYERS, R. E. Interocular transfer of pattern discrimination in cats following section of crossed optic fibers. *Journal of Comparative and Physiological Psychology*, 1955, **48**, 470-473.
- MYERS, R. E. Functions of corpus callosum in interocular transfer. *Brain*, 1956, **79**, 358-363.
- MYERS, R. E. Corpus callosum and visual gnosis. In J. F. Delafresnaye (Ed.), *Brain mechanisms and learning*. Oxford: Blackwell, 1961, Pp. 481-503.
- MYERS, R. E. Visual deficits after lesions of brain stem tegmentum in cat. *Archives of Neurology*, 1964, **11**, 73-90.
- MYERS, R. E., & SPERRY, R. W. Interhemispheric communication through the corpus callosum. *Archives of Neurology and Psychiatry*, 1958, **80**, 298-303.
- MYERS, R. E., SPERRY, R. W., & MCCURDY, N. M. Neural mechanisms in visual guidance of limb movement. *Archives of Neurology*, 1962, **7**, 195-202.
- NAUTA, W. J. H., & BUCHER, V. M. Efferent connections of the striate cortex in the albino rat. *Journal of Comparative Neurology*, 1954, **100**, 257-286.
- PAPEZ, J. W. The thalamic nuclei of the nine-banded armadillo (*tatusia novemcincta*). *Journal of Comparative Neurology*, 1932, **56**, 49-103.
- PAPEZ, J. W., & FREEMAN, G. L. Superior colliculi and their fiber connections in the rat. *Journal of Comparative Neurology*, 1930, **51**, 409-440.
- PEARCE, G. W. Some cortical projections to the midbrain reticular formation. In D. B. Tower & J. P. Schade (Eds.), *Structure and function of the cerebral cortex*. London: Elsevier, 1960. Pp. 131-137.
- PENFIELD, W. Epileptic automatism and the centrencephalic system. *Research Publications in the Association of Nervous and Mental Disorders*, 1952, **30**, 513-528.
- PENFIELD, W. Mechanisms of voluntary movement. *Brain*, 1954, **77**, 1-17. (a)

- PENFIELD, W. Some observations on the functional organization of the human brain. *Proceedings of the American Philosophical Society*, 1954, 98, 293-297. (b)
- PENFIELD, W. Studies of the cerebral cortex of man: A review and an interpretation. In J. F. Delafresnaye (Ed.), *Brain mechanisms and consciousness*. Springfield, Ill.: Charles C Thomas, 1954. Pp. 284-304. (c)
- PENFIELD, W. Consciousness and centre-encephalic organization. *First International Congress on the Neurological Sciences*. Brussels: 1957. Pp. 7-18.
- PENFIELD, W. *The excitable cortex in conscious man*. Springfield, Ill.: Charles C Thomas, 1958.
- PENFIELD, W. The interpretive cortex. *Science*, 1959, 129, 1719-1725.
- PENFIELD, W. Speech and perception. *Rehabilitation Monographs*, 1963, 23, 23-48.
- PENFIELD, W., & JASPER, H. *Epilepsy and the functional anatomy of the human brain*. Boston: Little, Brown, 1954.
- RANSON, S. W., & MAGOUN, H. W. The central path of the pupilloconstrictor reflex in response to light. *Archives of Neurology and Psychiatry*, 1933, 30, 1193-1204.
- RICH, I., & THOMPSON, R. Role of the hippocampo-septal system, thalamus, and hypothalamus in avoidance conditioning. *Journal of Comparative and Physiological Psychology*, 1965, 59, 66-72.
- RUSSELL, G. V. Interrelationships within the limbic and centreencephalic systems. In D. E. Sheer (Ed.), *Electrical stimulation of the brain*. Austin: Univer. Texas, 1961. Pp. 167-181.
- RUSSELL, I. S., & OCHS, S. Localization of a memory trace in one cortical hemisphere and transfer to the other hemisphere. *Brain*, 1963, 86, 37-54.
- SCHWARTZ, A. S., & CLARK, G. Discrimination of intermittent photic stimulation in the rat without its striate cortex. *Journal of Comparative and Physiological Psychology*, 1957, 50, 468-471.
- SECHZER, J. A. Successful interocular transfer of pattern discrimination in "split-brain" cats with shock-avoidance motivation. *Journal of Comparative and Physiological Psychology*, 1964, 58, 76-83.
- SMITH, K. U. Visual discrimination in the cat: V. The postoperative effects of removal of the striate cortex upon intensity discrimination. *Journal of Genetic Psychology*, 1937, 51, 329-369.
- SMITH, K. U. Visual discrimination in the cat: VI. The relation between pattern vision and visual acuity and the optic projection centers of the nervous system. *Journal of Genetic Psychology*, 1938, 53, 251-272.
- SPERRY, R. W. Cerebral organization and behavior. *Science*, 1961, 133, 1749-1757.
- SPERRY, R. W., STAMM, J. S., & MINER, N. Rerelearning tests for interocular transfer following division of optic chiasma and corpus callosum in cats. *Journal of Comparative and Physiological Psychology*, 1956, 49, 529-533.
- SPRAGUE, J. M., CHAMBERS, W. W., & STELLAR, E. Attentive, affective, and adaptive behavior in the cat. *Science*, 1961, 133, 165-173.
- STAMM, J. S., & SPERRY, R. W. Function of corpus callosum in contralateral transfer of somesthetic discrimination in cats. *Journal of Comparative and Physiological Psychology*, 1957, 50, 138-143.
- THOMPSON, R. Learning in rats with extensive neocortical damage. *Science*, 1959, 129, 1223-1224.
- THOMPSON, R. The interpeduncular nucleus and avoidance conditioning in the rat. *Science*, 1960, 132, 1551-1553.
- THOMPSON, R. Cortical and subcortical structures mediating visual discrimination habits in the rat. *Boletín del Instituto de estudios Médicos y biológicos, Universidad nacional de México*, 1963, 21, 451-466. (a)
- THOMPSON, R. Thalamic structures critical for retention of an avoidance conditioned response in rats. *Journal of Comparative and Physiological Psychology*, 1963, 56, 261-267. (b)
- THOMPSON, R., & MASSOPUST, L. C. The effect of subcortical lesions on retention of a brightness discrimination in rats. *Journal of Comparative and Physiological Psychology*, 1960, 53, 488-496.
- THOMPSON, R., & RICH, I. Differential effects of posterior thalamic lesions on retention of various visual habits. *Journal of Comparative and Physiological Psychology*, 1963, 56, 60-65.
- THOMPSON, R., RICH, I., & LANGER, S. K. Lesion studies on the functional significance of the posterior thalamomesencephalic tract. *Journal of Comparative Neurology*, 1964, 123, 29-44.
- TSAT, C. The optic tracts and centers of the opossum, *Didelphis virginiana*. *Journal of Comparative Neurology*, 1925, 39, 173-216.
- VALVERDE, F. Reticular formation of the albino rat's brain stem Cytoarchitecture

- and corticofugal connections. *Journal of Comparative Neurology*, 1962, 119, 25-53.
- VONEIDA, T. J. Performance of a visual conditioned response in split-brain cats. *Experimental Neurology*, 1963, 8, 493-504.
- WALLER, W. H. Progression movements elicited by subthalamic stimulation. *Journal of Neurophysiology*, 1940, 3, 300-307.
- WING, K. G., & SMITH, K. U. The role of the optic cortex in the dog in the determination of the functional properties of conditioned reactions to light. *Journal of Experimental Psychology*, 1942, 31, 478-496.
- WOODS, J. W. Behavior of chronic decerebrate rats. *Journal of Neurophysiology*, 1964, 27, 635-644.

(Received August 10, 1964)

original choice of the term "simplex" to describe the matrix. Jones argues that this interpretation is preferable in this case to the usual factorial interpretation because it enables description in terms of a single dimension (simplicity), whereas routine factor analysis would require several dimensions. He describes this approach as "molar," in contrast to the "molecular" approach of conventional factor analysis.

But there are difficulties with Jones' approach. In the first place, the same correlation matrix may be obtained from the converse equations describing decreasing simplification:

$$\begin{aligned}x_1 &= f_1 \\x_2 &= f_1 + f_2 \\x_3 &= f_1 + f_2 + f_3 \\&\dots \dots \dots [3] \\x_n &= f_1 + f_2 + f_3 + \dots + f_n\end{aligned}$$

Jones' choice of Equations 2 rather than Equations 3 to describe practice is based on somewhat imprecise psychological considerations; there is nothing in the correlational data themselves to indicate which is to be preferred.

Moreover, Jones appears to have based his defense of Equations 2 on the assumption that Equations 2 and 3 represent the only possible alternatives. This is not the case. We may suppose, for example, that increasing simplification occurs up to a point, when decreasing simplification begins, as in the following equations:

$$\begin{aligned}x_1 &= f_1 + f_2 + f_3 + f_4 \\x_2 &= f_2 + f_3 + f_4 \\x_3 &= f_3 + f_4 \\x_4 &= f_4 \\x_5 &= f_5 + f_4 \\x_6 &= f_6 + f_5 + f_4 \\x_7 &= f_7 + f_6 + f_5 + f_4\end{aligned} [4]$$

Clearly, we may go even further and consider any successive combinations of increasing and decreasing simplification. Indeed, we may rotate axes so as to render meaningless the notion of simplification, either increasing or decreas-

ing, and still retain a factorial representation which generates a simplex.

The concept of simplification, in Jones' account, also appears to depend on the proposition that the f_i of Equations 2 are themselves meaningful factors. It is therefore a further weakness of his position that he was unable to identify the f_i he obtained from Fleishman's (1960) data. This might suggest in fact that some rotation of axes would lead to more meaningful interpretation—and destroy the concept of simplification. However, it might be argued that Jones' attempt to identify the f_i was ill advised and that Guttman (1954) in his development of simplex theory had in mind a dimension whose main justification was precisely that it did *not* require the f_i to be identified. Perhaps so, but it appears that Guttman nonetheless based his choice of the term "simplex" on a model such as is represented by Equations 2, which implicitly demand that the f_i have some meaning. In a theoretical sense, if not in a predictive sense, the dimension envisaged by Guttman remains obscure.

The above criticisms apply to the interpretation of the simplex offered by Jones. It is not disputed that the simplex may provide a close fit to the practice matrix, in fact Jones' demonstration that this may be universally so is of great interest. It appears, however, that a more general interpretation of the simplex is required before the full meaning of the demonstration can be evaluated.

THE SIMPLEX AS A RANDOM WALK

Equations 2, 3, and 4 have an important characteristic in common: Each variable differs from the preceding one by just one factor, or vector (perhaps a better term than "factor" in this case since the factor content of these vectors is not clear). Moreover the vectors by which each pair of adjacent variables differ are mutually uncorrelated. This relation may be expressed as follows:

$$x_{i+1} = x_i \pm f_{i+1} [5a]$$

With a little algebraic manipulation, this equation may be written in standard score form as follows:

$$z_{i+1} = r_{i,(i+1)} \cdot z_i \pm \sqrt{1 - r_{i,(i+1)}^2} \cdot \phi_{i+1} \quad [5b]$$

where $r_{i,(i+1)}$ is the correlation coefficient between i^{th} and $(i+1)^{\text{th}}$ variables, and ϕ_{i+1} is the standard score equivalent of f_{i+1} .

Assuming only that the relation between scores is linear, Equation 5b provides a general model for the simplex in some unspecified factor space. Geometrically, the entire structure may be generated as follows. The first variable is represented as a unit vector in a hyperspace. The second variable is located by rotating the first through an angle of $\cos^{-1} r_{12}$. The third is located by rotating the second in a plane at right angles to the plane formed by the first two variables, through an angle of $\cos^{-1} r_{23}$. In general, the $(i+1)^{\text{th}}$ variable may be located by rotating the i^{th} variable in a plane at right angles to the hyperplane formed by the first i variables, through an angle of $\cos^{-1} r_{i,(i+1)}$. The final result is an n -dimensional structure. It could have been generated by beginning with the final variable and working backwards or by beginning with any intermediate variable and working outwards.

The reader may gain a better intuitive grasp of the simplex structure by considering the following "random-table model" in preference to the geometric model. Any column in a table of random numbers may be chosen to represent the first variable. The second variable may be obtained by adding the next column to the first. Succeeding variables may be generated by adding a new column each time. Columns may be subtracted instead of added at any stage. Provided that the number of scores in each column is large enough, the matrix of intercorrelations between the variables should give a good approximation to the simplex. By weighting each column before addition (or subtraction) with some appropriate scalar multiplier, it should be possible to simulate any given simplex.

This structure may be described as a *random walk* (whether through factorial hyperspace or through random tables), the word "random" being used in the sense that every step (ϕ_i) is uncorrelated with every other. This random-walk relation exists regardless of how the reference factors are defined or placed. In this respect its principal characteristic is molar rather than molecular, in the sense that Jones has emphasized. Moreover, it exists for every solution of the simplex which conforms to the basic linear factor model, since the structure is unaffected by any rotations of axes. Hence it is not necessary to choose between Equations 2, 3, or 4 or any other possible variants of these equations; all of these sets of equations may be thought of as describing a random walk.

A factorial interpretation of a random walk must be virtually impossible, particularly in the present context where each variable is in fact a practice trial on some task. The vectors ϕ_i , though random at least in the sense of being mutually uncorrelated, cannot be dismissed as representing error variance, since we are considering the simplex in unattenuated form. To assign factorial meaning to these vectors would require at least as many factors as there are variables.³ When each variable is a practice trial, moreover, it should be possible theoretically to obtain any number of factors, however large, by simply increasing the number of practice trials to the required amount. No conventional factorial interpretation can make sense of these properties.

A molar view of the random-walk representation of practice permits the somewhat trite observation that the trials are ordered in sequence; this sequence, of course, reflects exactly the temporal sequence in which the trials were performed. Kaiser (1962) has demonstrated that the simplex can be

³ Actually only half this number are required if the minimum rank model is chosen to describe the simplex (DuBois, 1960). Humphreys (1960) has already argued, however, that no sense can be made of the simplex if the minimum rank model is chosen.

interpreted as a ratio scale, but if we discard the notion of simplicity (or complexity), the problem is to decide what this scale measures.

In general, then, neither the molecular approach of factor analysis nor the more molar approach suggested by Jones provides a satisfactory theoretical interpretation of the practice simplex. Rather than abandon the simplex altogether, however, the following sections will explore the possibility of interpretation in terms of a linear model which will be presented as an alternative to the usual factor model.

ALTERNATIVES TO THE FACTOR MODEL

For N individuals, n tests, and r factors the conventional linear factor model may be written in matrix form as follows:

$$Z = MF \quad [6a]$$

where Z is the $n \times N$ matrix of test scores, M is the $n \times r$ matrix of factor loadings, and F is the $r \times N$ matrix of factor scores.⁴ It is not crucial to the remainder of this section to specify whether this equation shall represent the total factor space or the common factor portion only.

Equation 6a may be rewritten thus:

$$\begin{bmatrix} z_1 \\ z_2 \\ \cdot \\ \cdot \\ z_i \\ \cdot \\ z_n \end{bmatrix} = \begin{bmatrix} m_1 F \\ m_2 F \\ \cdot \\ \cdot \\ m_i F \\ \cdot \\ m_n F \end{bmatrix} \quad [6b]$$

where each z_i is a $1 \times N$ vector containing the scores of the N individuals on Test i , and each m_i is a $1 \times r$ vector containing the factor loadings of Test i on the r factors.

Equation 6b makes plain an implicit assumption of the factor model: that the matrix of factor scores, F , remains the

same for each test. That is, it is assumed that each individual retains the same score on each factor throughout administration of the test battery. In the majority of factor studies there may be no reason to doubt this assumption. However, there are certain situations in which its validity must be open to serious question, specifically when there is the possibility of transfer, learning, or maturational effects occurring between tests. In such circumstances the following more general equation may be more appropriate:

$$\begin{bmatrix} z_1 \\ z_2 \\ \cdot \\ \cdot \\ z_i \\ \cdot \\ z_n \end{bmatrix} = \begin{bmatrix} m_1 F_1 \\ m_2 F_2 \\ \cdot \\ \cdot \\ m_i F_i \\ \cdot \\ m_n F_n \end{bmatrix} \quad [7]$$

where F_i is the matrix of factor scores at the time when Test i is administered, and the F_i are not necessarily the same for different values of i .

Although a model of this kind may eventually be necessary for a satisfactory factorial analysis of tests involving learning, transfer, or maturational changes within the test battery, it is proposed here to consider a special case only. In the case of practice or learning, when the tests are in fact successive trials on the same test or task, it seems a reasonable assumption to make that the factor *loadings* of this task do *not* change from trial to trial; that is, let us assume that $m_i = m$, say, for all i . Equation 7 may now be written:

$$\begin{bmatrix} z_1 \\ z_2 \\ \cdot \\ \cdot \\ z_i \\ \cdot \\ z_n \end{bmatrix} = \begin{bmatrix} m F_1 \\ m F_2 \\ \cdot \\ \cdot \\ m F_i \\ \cdot \\ m F_n \end{bmatrix} \quad [8]$$

It might be objected at this point that Equations 7 and 8 run counter to the principal of parsimony, since factors are hypothetical constructs and can

⁴ Note that, following Harman's (1960) standard text, the symbol F has been chosen here to represent the factor *score* matrix. F is also widely used to represent the matrix of factor loadings.

always be defined so as to conform to Equations 6a and 6b; that is, one certainly *can* factor analyze the simplex in terms of the usual factor model and obtain factors. Equations 7 and 8 may therefore seem superfluous. However, as has already been argued, factors obtained from the simplex and based on the usual factor model as represented by Equations 6a and 6b are uninterpretable. Further, it will be argued in the following section that much better sense can be made of the practice simplex, at least, if the model represented by Equation 8 is chosen in preference to that represented by 6b, although ultimately a more complex model such as is represented by 7 may give the more penetrating solution.

INTERPRETATION OF THE PRACTICE SIMPLEX

Let us suppose now that the random-walk relation inferred above exists, not between the factor *loadings* from trial to trial (i.e., between the m_i of Equation 6), but between the factor *scores* (i.e., between the F_i of Equation 8). That is, we assume that the factor scores for each individual are subject to small random changes from trial to trial. Actually, it might be expected of course that scores on a practice task would improve from trial to trial. Since correlational data only convey information about scores relative to their mean values, the present approach is not inconsistent with the notion that mean scores improve from trial to trial. We need only suppose that changes are random about some mean change; again "random" in the sense that every set of score changes is uncorrelated with every other set. This state of affairs may be represented by the following equation:

$$F_{i+1} = a_{i+1}F_i + \sqrt{1 - a_{i+1}^2} \cdot C_{i+1} \quad [9]$$

where C_{i+1} is the $r \times N$ matrix of random elements (change scores), and a_{i+1} is some scalar multiplier.

It can be shown that Equation 9 generates a simplex as follows: First,

the correlation between any two trials, i and j , is given by:

$$r_{ij} = 1/N(z_i z_j') \quad [10]$$

where z_j' is the transpose of z_j . From Equation 8 we have $z_i = mF_i$ and $z_j = mF_j$, so that Equation 10 becomes:

$$r_{ij} = 1/NmF_i F_j' m' \quad [11]$$

Now consider any three trials, i, j , and k , such that $i < j < k$. We may consider these three trials to be consecutive trials by disregarding those intervening between them, and from Equation 9 we may write the following equations connecting F_i, F_j , and F_k :

$$\begin{aligned} F_j &= a_j F_i + \sqrt{1 - a_j^2} \cdot C_j \\ F_k &= a_k F_j + \sqrt{1 - a_k^2} \cdot C_k \end{aligned} \quad [12]$$

where

$$F_i C_j' = F_i C_k' = F_j C_k' = C_j C_k' = 0 \quad \text{(the null matrix)} \quad [13]$$

It follows from Equations 11, 12, and 13 that the correlations between trials i, j , and k can be written:

$$\begin{aligned} r_{ij} &= a_j / N \cdot m F_i F_i' m' \\ r_{ik} &= a_j a_k / N \cdot m F_i F_i' m' \\ r_{jk} &= a_k / N \cdot m F_j F_j' m' \end{aligned} \quad [14]$$

These correlations satisfy Equation 1 which defines the simplex.

Several arguments may be advanced in favor of this interpretation of the practice simplex rather than that based on the usual factor model:

1. The apparently insurmountable difficulties of interpreting a random-walk relation between factor loadings are overcome since it is assumed in the present interpretation that the factor content remains constant from trial to trial. The individuals, not the task, change; a possibility which Humphreys (1960) has recognized. This notion has the further advantage that a test may still be "cataloged" with its factor content specified regardless of the state of practice among the individuals, whereas in the approach exemplified by Fleish-

man's work it would be necessary to qualify the factor content depending on how practiced the individuals were.

2. The notion of a random-walk relation between scores is not a new concept in learning. For example, Estes (1959) has proposed a random-walk model for choice behavior. Generally, the present model accords well with stochastic theories of learning and may even provide a useful technique for the confirmation and extension of such theories.

3. Several workers have failed to identify any kind of learning factor based on correlations between score changes and other criterion variables (Simrall, 1947; Woodrow, 1946). This is consistent with the present view that score changes in learning are random.

4. The present interpretation would imply that any factors derived in the usual way by factor analyzing a simplex would be artifacts. Such "artifacts" are generally easily recognizable. Typically, they show systematic changes in loadings from trial to trial, tapering from a maximum at some point in the sequence of trials. If two artifactors are extracted, as is commonly the case, they are generally rotated to symmetrical positions such that one has a maximum loading towards one end of the sequence, the other a maximum towards the opposite end. Because of the simplex "border effect" (Guttman, 1954), however, there is a drop in loadings for each end trial. Finally, as might be expected, artifactors are generally uninterpretable in terms of any reference battery of tests. All of the above features are illustrated by the two factors obtained by Fleishman (1960) in a study of pursuit rotor learning, identified (or unidentified, rather) as "RP Specific I" and "RP Specific II" and shown in Figure 2 of that report. The same features are apparent in many of the factors obtained by Fleishman and his associates from practice or learning data.

The above arguments are consistent, then, first with the view that score changes in practice are random (about some mean change) and second with

the view that these changes represent changes in factor scores, not factor loadings. It is possible, however, that the case has been overstated. At least some of Fleishman's data give fairly convincing evidence for genuine change in factor loadings with practice (e.g., Fleishman & Rich, 1963). Further, Manning and DuBois (1962), using more sophisticated techniques than earlier workers such as Woodrow (1946), who were not able to do so, have reported evidence for a learning factor. Their finding suggests that score changes with learning or practice may not be altogether random; the walk may not always be as random as the rigorously defined simplex implies. We might speculate, perhaps, that practice is a stochastic process from the point of view of each individual's performance but that there may exist consistent differences between individuals in the parameters of this stochastic process. The simplex component of practice matrices might therefore reflect the stochastic properties of change *within* individuals, while what evidence there is for genuine change in factor content with practice might reflect parametric differences *between* individuals. Such differences might be expected to be difficult to isolate since they are confounded both by the stochastic properties of change within individuals and by measurement error.

SOME METHODOLOGICAL IMPLICATIONS

The issues discussed in this paper raise a number of methodological points:

1. The most important is that the practice matrix should not be factor analyzed or included in any matrix to be factor analyzed since it is likely that artifactors resulting from changes in factor scores rather than changes in factor loadings will appear. Determination of factorial change with practice should be based on correlations of trial scores with criterion factors or tests defined independently of the practice matrix itself. This requirement is satisfied in the study of two-hand coordina-

tion by Fleishman and Rich (1963) but not in Fleishman's (1960) study of rotary pursuit learning.

2. It is also important that determination of factor content at any stage in learning or practice be evaluated by reference to criterion tests administered *at that stage*. For example the random-walk interpretation of the practice simplex presented above would lead us to predict that correlations of trial scores with some criterion would in any case decrease with practice provided that the criterion task is administered before practice begins. This follows not because the practice task is becoming less like the criterion task but because the individuals themselves are becoming less like what they were when the criterion task was administered. Conversely, if the criterion task is administered after practice, we might predict increasing correlations.⁵ Fleishman and Rich (1963) do not state explicitly when in relation to practice their two criterion tests were administered, but if we assume them to have been administered before practice began, then the finding of increasing correlations between "kinesthetic sensitivity" and the practice task is more convincing demonstration of genuine factorial change associated with the test, not the individuals, than is the finding of decreasing correlations between "aerial orientation" and the practice task.

3. The confounding influence of changing factor scores might best be eliminated either by "removing" the simplex component using a technique such as Kaiser (1962) has suggested, or by carrying out separate factor analyses before and after practice. In the latter alternative the test batteries should include, on the one hand, the first trial and criterion tests administered before practice and, on the other hand, the final trial and the

same (or parallel) criterion tests administered after practice. Actually another issue, besides that involving changing factor loadings versus changing factor scores, is involved here. It concerns the nature of transfer. Supposing that scores on some factor represented in the practice task do change with practice, does it follow that the same changes will occur on this factor as it may be represented in other tasks? If the answer is yes, we would expect, on the assumption of changing factor scores but invariant factor loadings with practice, that factor analyses before and after practice would yield essentially the same results (Humphreys, 1960). If changes in factor scores on the practice task are *not* in general accompanied by changes in the same factor scores as measured by other tests (a possibility which would raise some new difficulties for factor theory), then we might expect that loadings of factors on the practice task should be smaller after practice than before but that the general pattern of loadings should remain the same. Finally, genuine change in factor content of the task should result in altered pattern of factor loadings after practice.

REFERENCES

- DuBois, P. H. An analysis of Guttman's simplex. *Psychometrika*, 1960, 25, 137-182.
- ESTES, W. K. A random walk model for choice behavior. In K. J. Arrow, S. Karlin, & P. Suppes (Eds.), *Mathematical methods in the social sciences*. Stanford: Stanford Univer. Press, 1959. Pp. 265-276.
- FLEISHMAN, E. A. A factor analysis of intra-task performance on two psychomotor tests. *Psychometrika*, 1953, 18, 44-55.
- FLEISHMAN, E. A. Abilities at different stages of practice in rotary pursuit performance. *Journal of Experimental Psychology*, 1960, 60, 162-171.
- FLEISHMAN, E. A., & HEMPEL, W. E., JR. The relation between abilities and improvement with practice in a visual discrimination task. *Journal of Experimental Psychology*, 1955, 49, 301-316.
- FLEISHMAN, E. A., & RICH, S. Role of kinesthetic and spatial-visual abilities in perceptual-motor learning. *Journal of Experimental Psychology*, 1963, 66, 6-11.
- GUTTMAN, L. A new approach to factor analysis: The Radex. In P. F. Lazarsfeld

⁵ Surprisingly, there seems to be a lack of evidence in the literature as to whether these predictions hold in fact. Authors either do not state when in relation to practice the criterion tasks were administered, or they confuse the issue by presenting the criterion tasks before practice to half the subjects and after practice to the other half.

- (Ed.), *Mathematical thinking in the social sciences*. Glencoe, Ill.: Free Press, 1954. Pp. 258-348.
- HARMAN, H. H. *Modern factor analysis*. Chicago: Univer. Chicago Press, 1960.
- HUMPHREYS, L. G. Investigations of the simplex. *Psychometrika*, 1960, 25, 313-323.
- JONES, M. B. Practice as a process of simplification. *Psychological Review*, 1962, 69, 274-294.
- KAISER, H. F. Scaling a simplex. *Psychometrika*, 1962, 27, 155-162.
- MANNING, W. H., & DuBOIS, P. H. Correlational methods in research on human learning. *Perceptual and Motor Skills*, 1962, 15, 287-321.
- REYNOLDS, B. Correlations between two psychomotor tasks as a function of distribution of practice on the first. *Journal of Experimental Psychology*, 1952, 43, 314-348.
- SIMRALL, D. V. Intelligence and the ability to learn. *Journal of Psychology*, 1947, 23, 27-43.
- WOODROW, H. The ability to learn. *Psychological Review*, 1946, 53, 147-148.

(Received July 23, 1964)

PSYCHOLOGICAL REVIEW

CATEGORY JUDGMENT: A RANGE-FREQUENCY MODEL¹

ALLEN PARDUCCI

University of California, Los Angeles

The range-frequency theory is concerned with category judgments, like "good" and "bad," or "large," "medium," and "small." A specific model derives the judgments from 2 basic assumptions: (a) The judge divides his psychological range into subranges whose relative sizes are independent of the stimulus conditions; and (b) he employs the alternative categories with equal frequency. The model uses judgments obtained when stimuli are presented with equal frequency to predict the judgments when stimuli are presented with unequal frequencies. These data are also used to evaluate the weighted-mean model for adaptation level. It is concluded that category judgments are more adequately explained by the range-frequency theory than by the theory of adaptation level.

This paper presents a theory of category judgment, tested in experiments with squares which vary in size. The experimental task requires the subject to judge each square as it is presented, using an ordered set of size categories, from "very small" to "very large."

In this task, as in everyday experience, the subject must establish his own standards for using the categories. "Large," like "beautiful" or "worthwhile," relates what is being judged to some implicit frame of reference. For the squares, the judgments adjust to the range of sizes in the experiment. After repeated presentations, "very large" is reserved for the largest, "very small" for the smallest, with the

other categories indicating the order of the intermediate sizes. However, the subject does more than merely rank the squares: His judgments reflect his sensitivity to the degree of difference between successive sizes and to their relative frequencies. His frame of reference or context for judgment appears to be the frequency distribution of stimulus values.

It is useful to study simple perceptual judgments because they are so dependent upon the immediate stimulus context. Since the presentations are controlled by the experimenter, the procedure permits independent manipulation of important features of the context. In the present experiments, the stimulus frequencies are manipulated in order to test a theory of judgment. The objective is to account for

¹ This research was supported by National Science Foundation Grant GB-1768 and Public Health Service Grant HD-00923.

the changes in judgment of each of the squares.

THE THEORY

The range-frequency theory treats category judgment as a compromise between two principles: One principle is concerned with how the judge divides the stimulus range, the other with how frequently he uses different categories.

The Range Principle

The first principle asserts that the judge uses the categories to subdivide the psychological range, the difference between the two extreme values of the stimuli that form the psychological context for judgment. A subrange corresponds to each category. The relative size of each subrange depends only upon the particular set of categories and is independent of stimulus conditions.

As a simple example, suppose there were just two categories, "large" and "small," which the subject used for stimuli in the upper and lower halves of his range. According to the range principle, a change in the stimulus frequencies would not affect his judgments of any of the stimuli unless it also changed his psychological range. Each category would always correspond to half the range.

Although dependent upon the physical range, the psychological range cannot be directly inferred from it. Thus, a subject may believe that the stimulus he is judging is the smallest that has been presented, even though it has been preceded by even smaller stimuli. His psychological range must be inferred from his judgments.

The Frequency Principle

The second principle asserts that the judge uses each category for a fixed proportion of his judgments. These

proportions may vary for different sets of categories, but we can often assume equal proportions. This will be done for the present tests of the theory. The frequency proportions are fixed in the sense that they are independent of the stimulus conditions, as with the relative sizes of the subranges.

But unlike the range principle, the frequency principle implies that the judgments are affected by the stimulus frequencies. Suppose that in the example in which "large" and "small" corresponded to equal subranges, the larger stimuli were presented more frequently than the smaller ones. "Large" would then be used more frequently. To conform to the frequency principle, "large" would have to correspond to a narrower subrange. In this case, the range and frequency principles would conflict.

The two principles conflict when stimuli from different parts of the range are presented with unequal frequencies. The theory postulates that the actual judgments reflect a compromise between the two principles. The nature of the compromise will now be specified in a model which yields quantitative predictions of the judgments.

THE MODEL

The range-frequency model deals with category limens. A limen is simply the boundary between two categories. In practice, it is defined as the stimulus value that, if it were presented, would be judged half the time with the higher of two categories (or with still higher categories). In this model, stimuli and limens are located on a Thurstone scale of equal discriminability (as described in Torgereson, 1958) rather than on a physical scale. Since the Thurstone scaling relates the judgments to the empirical limens, the model must predict these limens.

The model can be briefly stated, using three definitions and a single law of judgment:

D1. *Empirical limens*: the values of the boundaries between successive categories on a stimulus scale of equal discriminability.

D2. *Frequency limens*: the values the empirical limens would have if the categories were used with equal frequency.

D3. *Range limens*: the values the empirical limens would have if they divided the psychological range independently of the stimulus frequencies.

J4. Each empirical limen is the mean of the corresponding range and frequency limens.

This model simplifies the theory so that it can be tested experimentally. The compromise between the two principles of judgment places each empirical limen halfway between a range and a frequency limen. Thus, each pair of empirical and frequency limens determines a range limen. The range limens can be calculated from the judgments obtained when stimuli are presented with equal frequency. Since

the range limens are independent of the stimulus frequencies, they can be used to predict the empirical limens and judgments for stimuli presented with unequal frequencies.

Calculation of the Limens

This section describes how the different limens are obtained and their relationships to the actual judgments. The examples should make the meaning of the limens more concrete. However, this material is not essential to a general understanding of the range-frequency model, and the reader can follow the experimental tests and evaluation without working through the details of this section.

Figure 1 illustrates how the empirical limens are located by plotting the proportions of the judgments that place each of the stimuli above each of the empirical limens. Thus, the largest square, Stimulus 9, is shown as being placed in the topmost category (i.e., above the limen separating Categories 5 and 6) for about 50% of its presentations, and in either of the top two categories for about 95% of its presentations.

Figure 2 schematizes the location of the frequency limens for the same Rectangular

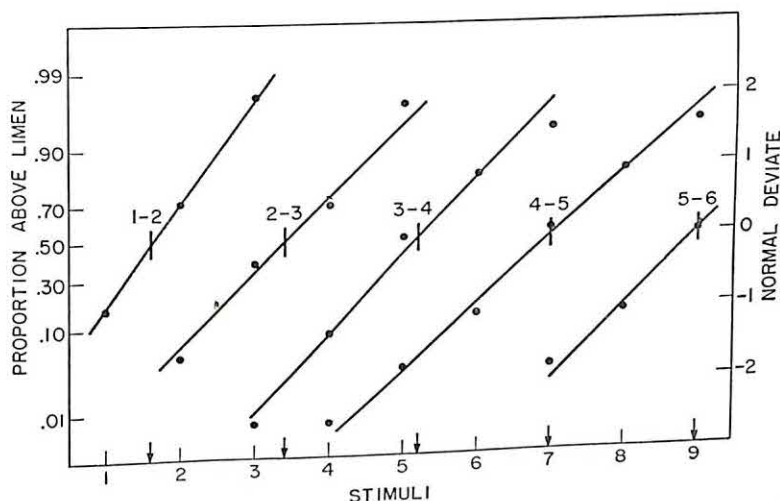


FIG. 1. Empirical limens determined from proportions of judgments greater than each limen for each stimulus in Rectangular distribution.

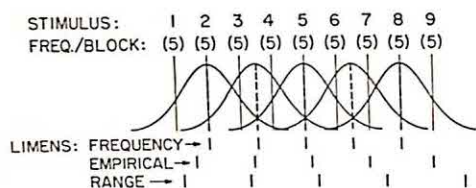


FIG. 2. Distributions of hypothetical values of frequency limens and their relationships to empirical and range limens for Rectangular distribution.

distribution of stimulus presentations that yielded the judgments for Figure 1. Variability of judgment is shown in Figure 2 as a normal distribution of values for each of the five frequency limens. This variability is attributed to shifting locations of the limens on the psychological scale of size. Since the variability is in the relationship between the stimuli and the limens, it could be portrayed equally well using a distribution of values for each of the nine stimuli or, following Thurstone, using distributions of values for the stimuli and also for the limens.

The SD of each of the normal distributions in Figure 2 is one stimulus step, the difference between successive stimulus values on the psychological scale of size. This is the approximate stimulus difference between the .50 and .84 values for each of the five functions in Figure 1, so that the simplest Thurstone scaling of these data would yield stimulus values separated by unit variability. The range-frequency model could also be applied by using different SDs for the different limens, by scaling the stimulus dimension with one of the more complex Thurstone procedures, or by a procedure for equalizing the differences between successive range limens.

The frequency limens in Figure 2 mark off equal frequencies of category use. Thus, the hypothetical momentary values contributing to the rightmost or highest frequency limen are above the value of Stimulus 9 for 1/6 of the presentations, above Stimulus 8 for 1/2 of the presentations, and above Stimulus 7 for 5/6 of the presentations. Since the nine stimuli are presented with equal frequency in the Rectangular distribution, the highest category would be used for $1/9(1/6 + 1/2 + 5/6) = 1/6$ of the total number of stimulus presentations.

Each of the five frequency limens, either for the Rectangular or for any other distribution, can be determined by successive ap-

proximation so as to mark off equal frequencies of category use. It is convenient to perform this approximation on normal-probability paper, starting with the empirical limens. Thus, the proportions indicated by the best-fit line drawn for the top limen in Figure 1 are approximately .02, .16, and .50. Cumulative multiplication of these proportions by the proportion of trials on which the corresponding stimuli were presented (1/9 for the Rectangular distribution) yields .076 as the value representing the proportion of presentations judged with the top category. This is less than the equal-frequency proportion which is the reciprocal of the number of categories (.167 for six categories). The frequency limen must therefore be lower than the empirical limen. To locate the top frequency limen, one shifts the best-fit line for the top empirical limen to the left, without changing its slope. The proportions of the stimulus presentations are then noted for the new location. If the cumulated proportion of all presentations is again less than the equal-frequency proportion, the line must be relocated still further to the left; if greater than the equal-frequency proportion, it must be shifted back to the right. For the Rectangular distribution, the line for the top frequency limen was finally located so that its .50 or limen value corresponded to Stimulus 8. Thus, the normal curve that represents the distribution of momentary values for the top frequency limen is centered over Stimulus 8 in Figure 2. The same procedure was applied to locate the other frequency limens.

The model asserts that each of the empirical limens is the mean of the associated frequency and range limens. Once the empirical and frequency limens have been determined, the range limens can be computed algebraically. Since the top empirical limen is 9 and the top frequency limen is 8 for the Rectangular distribution, the corresponding range limen must be 10.

The relationship between the different limens is shown at the bottom of Figure 2. In this example, only the frequency limens are symmetrical with respect to the stimulus values. The deviation from symmetry is toward higher values for the range and the empirical limens. By analogy to the positive time-order error found for comparative judgments of visual size, one might speculate that the subjects remember the previous presentations as larger than they were, or that subjects compare the presented squares with either a larger set or with larger areas in the visual field (McClelland,

1943). Whatever the basis for this phenomenon, the range limens provide a quantitative index for the shifted frame of reference. This index is free of the postulated bias toward equal use of the different categories.

Frequency limens can be located for any distribution of these same nine stimuli, assuming the same SDs but different locations for each of the limens. Figure 3 shows the limens for an irregular distribution of presentation frequencies. Hypothetical normal curves are again located so as to mark off equal frequencies of category use. For this distribution, the proportion of the area cut off by each stimulus must be weighted for the relative frequency with which the stimulus is presented. The frequency limens were again located by successive approximation, each category including 1/6 of the presentations. Consequently, these limens are most closely spaced in that part of the range where the stimulus values occur most frequently.

The range limens are assumed to have the same values for any distribution of these nine stimuli. Each range limen for the Rectangular distribution can then be averaged with the corresponding frequency limen for any other distribution to predict the empirical limen for the other distribution. These predicted values can then be com-

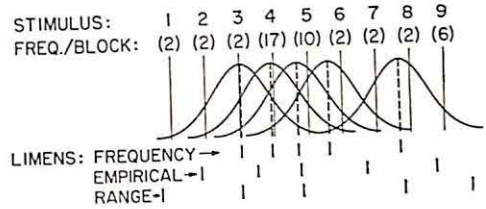


FIG. 3. Distributions of hypothetical values of frequency limens and their relationships to empirical and range limens for Low-Median distribution.

pared with the empirical limens. Or instead, as in Figures 2 and 3, the empirical limens can be combined with the frequency limens to calculate the range limens. According to the model, the latter should have similar values for different distributions of the same stimuli.

The stimulus-judgment matrix can be predicted for each distribution using the predicted empirical limens and the slopes fitted for the Rectangular distribution. The obtained frequencies can then be compared with these predicted frequencies. However, it is easier to evaluate a comparison based upon the predicted mean judgments. These are readily calculated from the predicted stimulus-judgment matrix.

TABLE 1
DISTRIBUTIONS OF STIMULUS PRESENTATIONS

Stimulus number	1	2	3	4	5	6	7	8	9
Width (in cm.)	5.4	6.7	8.3	10.5	13.3	16.2	18.7	20.7	23.2
Distribution	Stimulus frequencies								
Rectangular	5	5	5	5	5	5	5	5	5
Negatively Skewed	2	2	3	4	5	6	7	8	8
Positively Skewed	8	8	7	6	5	4	3	2	2
Normal	2	3	5	8	9	8	5	3	2
U-Shaped	2	3	5	8	9	8	5	3	2
Low-Median	8	6	4	3	3	3	4	6	8
High-Median	2	2	2	17	10	2	2	2	2
Low-Midpoint	6	2	2	2	10	17	12	0	0
High-Midpoint	2	2	2	12	9	6	2	2	2
Low-Midpoint-Median	0	0	12	6	9	12	2	2	0
High-Midpoint-Median	2	2	5	13	5	5	7	6	2
Low-Mean	0	6	7	5	5	13	5	3	3
High-Mean	11	3	3	3	5	11	3	3	11

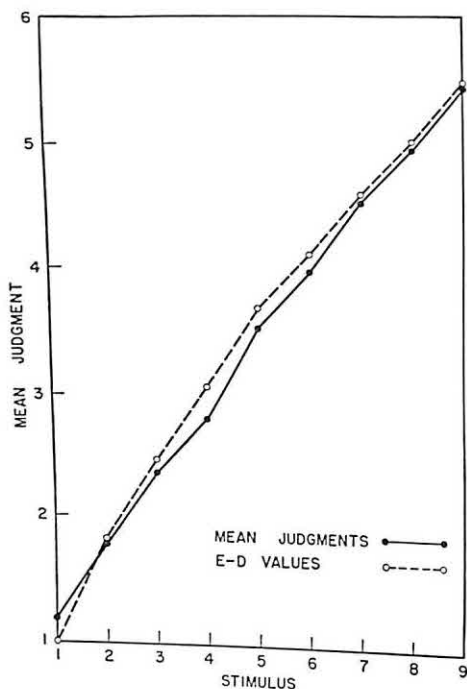


FIG. 4. Mean judgments and Thurstone equal-discriminability values for stimuli in Rectangular distribution.

EXPERIMENTAL TESTS OF THE MODEL

New Data

Five different distributions of stimulus presentations were used for the same set of nine squares. Frequencies per block of 45 presentations are shown in the top half of Table 1. Each distribution was presented to four different subgroups of eight subjects each, the random order of presentation varying for the different subgroups. Presentation was in two blocks of 45. Only the second block was tabulated so that attention is restricted here to judgments made after at least one exposure to the entire distribution. The subject's task was to judge each square in comparison with the other squares in the series. Judgments were made with the following six categories: "very small—1," "small—2," "slightly smaller than average—3," "slightly

larger than average—4," "large—5," and "very large—6." Additional details of procedure have been presented elsewhere (Parducci, 1963).

Figure 4 shows that the mean judgments are linearly related to the rank-order values of the stimuli when the latter are presented with equal frequency. The plot for the Thurstone values (calculated assuming equal stimulus and category dispersions) is more negatively accelerated. However, both functions suggest that the stimuli are spaced approximately equally on the psychological scale of size, as was assumed for the plots in Figures 1, 2, and 3.

Predicted values were calculated for the mean judgments for the Negatively Skewed, Positively Skewed, Normal, and U-Shaped distributions in Table 1. Frequency limens were calculated separately for each of these four distributions. Range limens were determined using the data from the Rectangular condition, as in Figure 2. The predicted value for each empirical limen is the mean of each pair of frequency and range limens. These means were then used to construct the stimulus-judgment matrices from which the predicted mean judgments were computed. This procedure bases the predictions solely upon the judgments of the Rectangular distribution so that all predictions could be made before any other distributions were presented for judgment.

The general trend of the judgments in Figures 5 and 6 is successfully predicted by the model. The predicted functions capture the overall difference in the level of judgment for the Skewed distributions in Figure 5 and the difference in slope for the Normal and U-Shaped distributions in Figure 6. The model also predicts the point of crossover for the two functions in Figure 6.

Nevertheless, some of the errors of prediction are systematic. The standard error for each empirical point, based upon the variability between the mean judgments of the same stimulus by the different subjects, is on the order of .1 category step. However, this overestimates the significance of the errors of prediction since the predicted values are themselves based upon variable data from the Rectangular distribution, and the different errors are not independent.

Old Data

The bottom eight distributions in Table 1 were studied 2 years earlier, and some of the mean judgments were reported before the present model was developed (Parducci, 1963). Figures 7-10 show the predicted functions along with the actual mean judgments.

The model again captures the general relationship between the sets of means obtained for each pair of distributions. But again, some of the differences between predicted and obtained values leave room for improvement.

Minor changes in procedure were required for the derivation of these predictions. Although the experimental procedure was very similar, the judgments had been collected in a different laboratory. Only the Rectangular distribution was used in both laboratories. The category limens were higher for the earlier data using this distribution, and the mean judgments were up to $1/3$ category step lower. Consequently, the prediction for the distributions described in this section are based on the earlier data for the Rectangular distribution. Figures 1 and 2 were also based on these data.

These eight distributions had been selected to test an earlier formulation of the range-frequency theory that predicted differences for the middle limen when either the medians or midpoints of the stimulus distributions

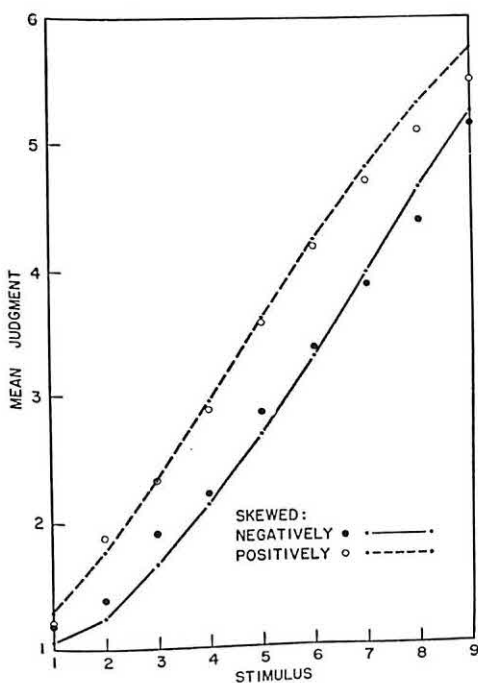


FIG. 5.

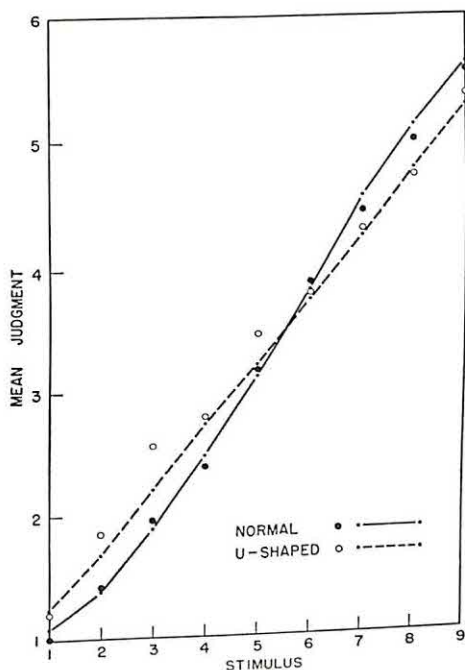


FIG. 6.

FIGS. 5-6. Mean judgments for new distributions of stimuli, predicted functions, and empirical points.

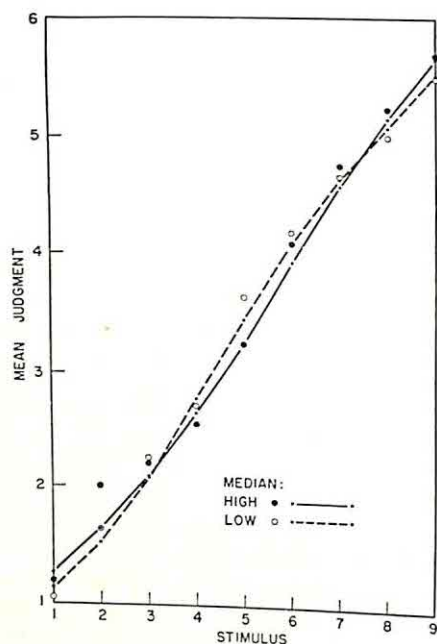


FIG. 7.

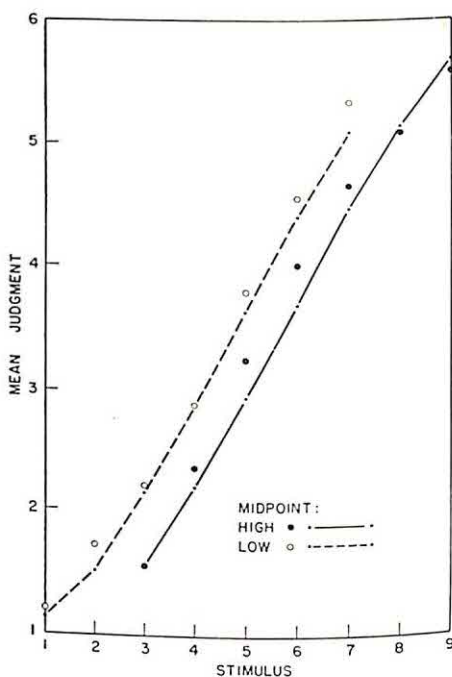


FIG. 8.

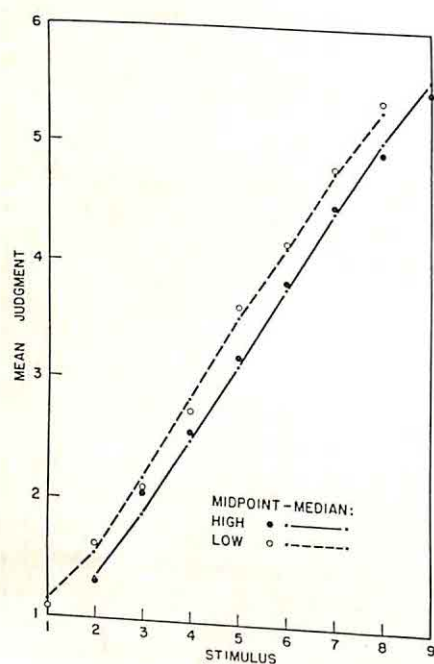


FIG. 9.

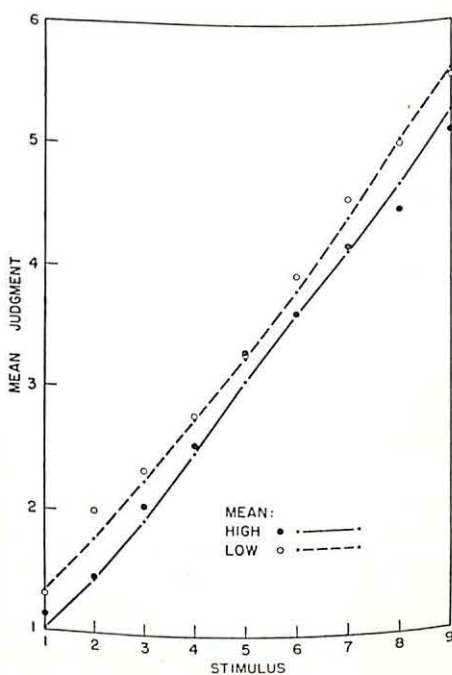


FIG. 10.

FIGS. 7-10. Mean judgments for old distributions of stimuli, predicted functions and empirical points.

were different. The stimulus means are identical for all but the last two of these distributions. The medians differ by one stimulus step for the Low- and High-Median distributions, but both have the same midpoint (the mean of the two end values). The medians are identical but the midpoints vary for the Low- and High-Midpoint distributions, and both measures vary together for the two Midpoint-Median distributions. Since the stimulus range was reduced by a factor of 7/9 for the Midpoint distributions and 8/9 for the Midpoint-Median distributions, the hypothetical range limens were compressed by the same factors in order to make predictions from the model. For each truncated distribution, the lowest range limen was set slightly above the lowest stimulus (as in Figure 2), the same factor being applied to determine this difference and also each of the successive subranges.

Only the stimulus means vary for the Low- and High-Mean distributions which were judged under somewhat modified instructions. Subjects were required to record each presentation number, 1 through 90, in the column labeled with the appropriate category on their response sheets (Parducci, 1963). Column recording raised the empirical limens so that the predictions for these two distributions were based on data from a separate group which had judged the Rectangular distribution under the modified instructions. All predictions were calculated in accordance with the procedure used for the new data.

Comparisons of Range Limens

Empirical limens were computed from the judgments of each distribution following the procedure described for Figure 1. These were combined with the frequency limens to calculate separate sets of range limens for each of the 13 distributions in Table 1. These range limens have similar values for all distributions with the same stimulus range, but they span a narrower range for the distributions whose stimulus ranges are narrower. The differences between successive, empirically determined range limens, that is, the subranges, are proportional to the stimulus range. In general, the five limens in each set are equally

spaced. This suggests that the model might be simplified further by incorporating both equal-subrange and equal-frequency principles.

Weighting the Limens

Taking the unweighted mean of each pair of range and frequency limens is an arbitrary simplification. However, moderate changes in weighting would scarcely affect the predictions, and the errors of prediction were systematically biased toward neither the range nor frequency limens. A separate empirical weighting was determined algebraically for each frequency limen for each distribution, using the associated empirical and frequency limens and the range limen from the Rectangular distribution. The mean weighting of the frequency limen is .53 with a standard error of .05. The weighting might be expected to vary systematically with manipulation of other conditions such as the discriminability of the stimuli or the instructions for judgment.

EVALUATION

It is clear that the distribution of stimulus frequencies has important effects upon judgment. It is also clear that the range-frequency model describes both the direction and the general order of magnitude of these effects. How well does it compare with alternative interpretations of judgment?

Adaptation-Level Theory

The most ambitious and influential approach to the problem of predicting the judgments of each stimulus has been Helson's theory of adaptation level (AL). The theory has been used to describe a wide variety of situations (Helson, 1964). Its first applications were to category judgments (Helson, 1938, 1947), and it is with respect to

the present type of data that the theory has had its most rigorous tests.

The basic assumption of AL theory is that the psychological effects of the stimulus context can be represented by a single stimulus value, AL. This is the value of the stimulus judged "average" (3.5 in the present task). The AL model that has been applied to psychophysical data (Helson, 1964, pp. 197-202) asserts that judgments are linearly related to the difference between each of the stimuli being judged and AL. This assumes that the stimuli have been suitably scaled with respect to the psychological dimension of judgment. The linearity of the judgment function for the Rectangular distribution (Figure 4) indicates that these squares have been suitably scaled. AL theory predicts that the other functions will be just as linear as the Rectangular function. To the degree that they are, AL theory predicts the present data.

Although the empirical functions are all roughly linear, straight-line fits would miss their other characteristic features. Thus, the functions intertwine for the Low- and High-Median distributions (Figure 7). The same intertwining has been found using different methods of recording judgments and also for such varying dimensions as lifted weight and numerosness of dots (Parducci, 1963). Linear fits would also miss the pinching at the middle stimulus for the Low- and High-Mean distributions (Figure 10).

A second basis for comparison is provided by the differences in the value of AL for different distributions of stimuli. The predictions for these differences are derived from the pooling feature of the theory which asserts that AL is a weighted mean of all of the stimuli affecting judgment. It can be assumed that the effects of

earlier stimulation and also of background stimuli do not differ systematically for the present experimental conditions. In this case, the simple AL model, the one that is applied to this type of situation (Helson, 1964, p. 132), describes AL as a linear function of the mean of the distribution of stimuli presented for judgment. With the exception of the Negatively and Positively Skewed and the Low- and High-Mean distributions, the present distributions were selected so that their means were each 5.0 on the stimulus scale. Contrary to the AL model, the empirical ALs, that is, the stimuli judged 3.5, have the same value for two of the distributions with varying means (Figure 10) but differ systematically for distributions with the same mean (Figures 7, 8, and 9). By contrast, the range-frequency model correctly predicts both the direction and general order of magnitude of the differences in AL.

Finally, a third basis for comparison is provided by the slopes of the judgment functions. According to the AL model, slope is determined by the number of judgment categories and by other factors which were not manipulated in the present research (Helson, 1964, pp. 199-201). Thus, the AL model predicts that the slopes of the functions in Figures 4 through 10 should all be equal. They are not. The slope is greater for the Normal than for the U-Shaped distribution (Figure 6) and for the distributions with narrower ranges (Figures 8 and 9) than for those with wider ranges (Figures 4, 5, 6, 7, and 10). Again, the range-frequency model does much better, in this case capturing the linear component of each of these functions.

With respect to each of the above three bases for comparison, the AL model proves less adequate than the range-frequency model.

Could another model be developed which would handle the present data and yet remain faithful to the general approach of AL theory? One possibility is that the simple weighted-mean model neglects important stimulus differences. As applied to other situations, AL theory weights each of the series stimuli in direct proportion to its frequency of presentation. But the background, anchors, and other stimuli given special attention in the task for judgment are weighted in accordance with their observed effects upon AL. If greater weighting were given the extreme values of the regular series, the smallest and largest squares in the present distributions, the AL predictions would be greatly improved. This change in the model would handle the effects of range. It might also be possible to find some system for weighting the stimuli which would account for the frequency effect, but then the scale of judgment could not be characterized by the single value, AL.

If the simple weighted-mean model for AL does so poorly with the present data, how has it proved useful in a wide variety of other situations? Were the present experiments different in crucial ways? The shifts in AL, shown here for squares, have also been reported for lifted weights and patterns of dots (Parducci, 1963), for lines which vary in length (Parducci & Marshall, 1961a), and for numerals presented in ordered arrays or in random succession (Parducci, Calfee, Marshall, & Davidson, 1960; Parducci & Marshall, 1961b). The general forms of the present judgment functions were also obtained with the other dimensions, indicating that they are relatively independent of the specific stimulus dimension.

The contrast between the previous success and the present failure of the AL model may simply reflect differ-

ences in the precision of the data used to test the model. As reported by Helson (1964), "groups judging identical stimuli may differ, on the average, by as much as a whole category step [p. 136]." And the classic study of distribution effects (Johnson, 1944) used only a single subject for each distribution. None of the studies cited as consistent with the AL model was sufficiently sensitive to demonstrate the deviations from the model reported here.

Although the present study questions the AL account of category judgment, it does not necessarily contradict the AL description of the effective stimulus for judgment. The prime value of the AL approach may be in specifying what is being judged. Thus, AL theory describes how the color of a sample changes with change in the reflectance of the immediate background, and it appears to provide a useful account of what happens to the memory for a standard when stimuli are interpolated between standard and comparison stimuli. The present data reflect only upon its adequacy as a theory of category judgment.

Conclusion

Consider the differences between the AL and the range-frequency interpretations of category judgment. AL theory characterizes the state of the organism with a single value, AL, which varies with the stimulus context. Like the mean of a set of scores, AL is an average of different stimulus effects. It represents the pooled effect of all the stimuli forming the context for judgment. Each judgment is determined by the relationship between the judged stimulus and AL. Therefore, the entire scale of judgment is determined by AL. Although attractive in its simplicity, this formulation cannot account for the evidence

that judgments may vary markedly even though there is no difference in AL (Figures 6 and 10) or that there may be differences in AL even though specific stimuli evoke the same judgments (Figures 5 and 7). The theory of AL cannot account for differences in the form of the judgment functions.

The range-frequency theory provides a more detailed characterization of the relationship between judgment and different features of the stimulus context. Special emphasis is given to the psychological representation of the stimulus range, the effective stimulus extremes. Although these clearly depend upon the extreme values of the physical distribution, further research is needed to clarify this dependency. The range-frequency theory also emphasizes the frequencies with which different categories are used. The present model assumes that there is a response tendency toward equal use of the judgment categories so that prediction of the judgments depends in part upon the frequencies with which stimuli are presented from each part of the range. Instead of comparing each stimulus with a single value, AL, the subject compares each stimulus with a set of category limens. This set is itself an average of two other sets, the hypothetical range and frequency limens. The subject could not describe these hypothetical limens, but his judgments are made as though he were compromising between them.

The advantage of the range-frequency approach is that it describes the effects of variation in stimulus frequency for the most reliable data available. These contextual effects il-

lustrate the nature of category judgment, and that is what the theory of adaptation level was designed to explain. The relative success of the present model indicates that the range-frequency theory provides a more adequate account of category judgment.

REFERENCES

- HELSON, H. Fundamental problems in color vision: I. The principle governing changes in hue, saturation, and lightness of non-selective samples in chromatic illumination. *Journal of Experimental Psychology*, 1938, 23, 439-476.
- HELSON, H. Adaptation-level as frame of reference for prediction of psychophysical data. *American Journal of Psychology*, 1947, 60, 1-29.
- HELSON, H. *Adaptation-level theory*. New York: Harper & Row, 1964.
- JOHNSON, D. M. Generalization of a scale of values by the averaging of practice effects. *Journal of Experimental Psychology*, 1944, 34, 425-436.
- MCCLELLAND, D. C. Factors influencing time error in visual extent. *Journal of Experimental Psychology*, 1943, 33, 81-95.
- PARDUCCI, A. Range-frequency compromise in judgment. *Psychological Monographs*, 1963, 77(2, Whole No. 565).
- PARDUCCI, A., CALFEE, R. C., MARSHALL, L. M., & DAVIDSON, L. P. Context effects in judgment: Adaptation level as a function of the mean, midpoint, and median of stimuli. *Journal of Experimental Psychology*, 1960, 60, 65-77.
- PARDUCCI, A., & MARSHALL, L. M. Context effects in judgments of length. *American Journal of Psychology*, 1961, 74, 576-583. (a)
- PARDUCCI, A., & MARSHALL, L. M. Supplementary report: The effects of the mean, midpoint, and median upon adaptation level in judgment. *Journal of Experimental Psychology*, 1961, 61, 261-262. (b)
- TORGERSON, W. S. *Theory and methods of scaling*. New York: Wiley, 1958.

(Received September 22, 1964)

PERCEPTUAL ADAPTATION TO INVERTED, REVERSED, AND DISPLACED VISION¹

CHARLES SAMUEL HARRIS

University of Pennsylvania

Recent research has shown that a simple form of adaptation to prism-produced displacement of the visual field consists primarily of a proprioceptive change—a change in the felt position of the arm seen through prisms—rather than a visual, motor, or visuomotor change. More complex sorts of adaptation (to inversion, reversal, and other optical transformations) can also be understood as resulting from changes in the felt locations of parts of the body relative to other parts. Contrary to the usual empiricist assumption, vision seems to be very stable, whereas the position sense is remarkably flexible. When the 2 senses provide discrepant information, it is the position sense that changes.

For over a century, psychologists have been experimenting with optical devices that displace, reverse, or invert the retinal image. When a person first puts on such a device, he misses things he reaches for and bumps into things he is trying to walk around. But after a while he adapts. He ends up behaving normally despite the optical distortion.

Typically, experimenters have accepted this adaptation as evidence for or against various theories about the origin of visual space perception in the infant. But even if one hesitates to generalize from adult behavior to infant development, adaptation to optical distortions is of interest in revealing how perceptual-motor systems work and how they can be modified.

¹ Based in part on a doctoral dissertation submitted to the Department of Psychology, Harvard University. Preparation of this paper was supported by NIMH Grant MH-10,711 and NSF Grant GB-3546. Some of the research cited was supported by NSF and NIMH predoctoral fellowships and by an NSF postdoctoral fellowship. I am grateful to Charles R. Hamilton, Judith R. Harris, Richard Held, Alice Isen, R. Duncan Luce, Jacob Nachmias, Fred Stollnitz, and Benjamin W. White for their helpful criticisms and suggestions.

Recently there has been much concern with the mechanisms for adapting to optical distortions and with the conditions that are necessary for such adaptation to take place. Less attention has been given to the *end product* of adaptation. What change does the adaptation procedure produce in the subject? How does the adapted subject differ from one who has not adapted?

Previous investigators have offered diverse answers to this question. For example, Kohler (1964) and Taylor (1962) believe that adaptation results in a change in visual perception. Smith and Smith (1962), on the other hand, claim that it consists mainly of learning specific motor responses. Held and Freedman (1963) say that adaptation "represents a change in state of the relevant sensorimotor control system" based on the storage of "newly correlated information" derived from "the one-to-one relation between movement and its sensory feedback [p. 457]."

This paper proposes another interpretation of adaptation: that it consists of changes in the position sense for various parts of the body. A

change in position sense has been clearly demonstrated in one form of adaptation to displaced vision. The extension of this interpretation to other forms of adaptation is more speculative but seems to make sense out of a mass of otherwise perplexing data. (For summaries of earlier experimental work on adaptation see Held & Freedman, 1963; Kohler, 1964; Smith & Smith, 1962; Taylor, 1962.)

The Position Sense

Even in the dark we can perceive the relative locations of the various parts of our bodies. The sense that enables us to do this will be referred to as the *position sense*, and the perception of the position will be called a *felt position*. Changes in the position sense will be called, for want of a better adjective, *proprioceptive* changes. (The term *kinesthesia* will be restricted to the perception of movements of parts of the body.)

The position sense is a psychological phenomenon; its physiological basis has not yet been conclusively established. Receptors in the joints seem to play the major role (Rose & Mountcastle, 1960); however, efferent activity may enhance the responses of these receptors, making the position sense more accurate during active movement (Lloyd & Caldwell, 1965). The fact that monkeys can perform acts with a deafferented limb (Taub, Elman, & Berman, 1964) suggests that a "sense of innervation," registering the motor outflow to the limb, may be able to take over the functions of sensory inflow. Indeed, motor outflow seems to be the sole basis for registering the position of the eyes (Brindley & Merton, 1960; Helmholtz, 1962b; Ludvig, 1952). If motor signals do play this role, though, the nervous system must somehow register the *positions*

called for, not the *movements*; otherwise we would lose track of body parts whenever they were moved (or kept from moving) by an outside force.

Although information registered by the position sense is usually available to introspection, we are not constantly aware of the locations of all of our body parts. And sometimes a subject's conscious perception of the positions of some body parts (especially his eyes) is vague and variable, even though there is abundant behavioral evidence that these positions are being "taken into account" accurately. In general, the hypotheses presented in this paper apply whether the position information is conscious, potentially conscious, or not available to consciousness.

ADAPTATION TO DISPLACED RETINAL IMAGES

Arm Adaptation

Adaptation to inversion or reversal of the visual field may take many days or even weeks. However, as Helmholtz reported in 1866, a person can adapt to sideways displacement of the visual field in just a few minutes (Helmholtz, 1962b, p. 246).

If you look through prisms that displace the apparent locations of seen objects to the right, for example, and try to reach quickly for something, you will miss it by reaching too far to the right. But after just a few more attempts, your aim will improve considerably. When the prisms are then removed, however, you will reach too far to the left. For convenience, both the improved reaching while wearing prisms and the aftereffect when they are removed will be referred to as *adaptation* (i.e., adjustment to new conditions), since they are presumably manifestations of a single underlying change. The amount of adaptation (the *adaptive shift*) is indicated by the difference between a subject's

responses on pre- and postadaptation tests. (During these tests the subject must not be allowed to see his hand; otherwise, by moving it slowly and guiding it visually, he would always be able to point correctly.)

Proprioceptive Changes. If a person's eyes are closed when he first puts on displacing prisms, he is surprised when he opens his eyes and looks at his hand. Because the prisms shift its visual image, his hand does not appear to be where he felt it was. If the discrepancy between the seen and felt locations of the hand is to be eliminated, either the person's visual perception or his position sense (or both) must shift.

According to the proprioceptive-change hypothesis, the subject comes to feel that his arm is where he saw it through prisms—even though this makes that arm's position sense erroneous (nonveridical). That is, after such a change, the subject's judgment of that arm's position relative to any other part of the body will be incorrect. If the prisms are removed and the subject tries (without seeing his hand) to reach for a target that he sees in a certain place, he will move his hand until he feels that it is in that place—but it will actually be off to one side of it. The same thing will happen if he tries to point at a sound or simply to point straight ahead. Only when judging the whereabouts of his hand relative to objects seen through prisms will he be accurate.

It is not clear a priori whether a proprioceptive shift would make a subject misperceive arm positions other than those he saw while adapting. Since neurons in the proprioceptive system have rather large receptive angles (Mountcastle, Poggio, & Werner, 1963), a change in the operating level of proprioceptive neurons in the central nervous system might exert an

effect over a wide range of arm positions. At any rate, the presence or absence of such a shift should depend mainly on the actual position of the arm, not on the movements by which the position was reached.

Other Interpretations. Five other simple, plausible conceptions of the nature of adaptation can also account for the rapid improvement in reaching for objects seen through prisms. Each, however, suggests a different set of predictions about other behavior. These five conceptions, which are often implicit rather than explicit in previous investigators' writings, have been presented in greater detail elsewhere (Harris, 1963a). They are described briefly below, together with some of their predictions about a subject who adapts by pointing with one arm, using a stereotyped arm movement, at a single target seen through prisms.

1. *Conscious correction of one's aim.* When the subject misses the target, he realizes that the prisms are deceiving him about the target's location and so deliberately aims to one side of visual targets; when the prisms are removed, he goes back to pointing normally.

2. *Altered visual perception.* A changed translation from retinal image to perception makes a target which at first looked off to the side appear to be straight ahead. This new perception can be demonstrated by any appropriate judgment of, or response to, a visual target seen with or without prisms.

3. *Reorientation of the perceptual frame of reference.* Perception of all external stimuli, visual or auditory, is shifted to one side; perception of the arms, however, is unaffected (if perception of the arm shifted too, the

TABLE 1
TEST PERFORMANCE PREDICTED BY SIX INTERPRETATIONS
OF ADAPTATION TO DISPLACED VISION

Test task	Proprioceptive change in the arm	Conscious correction	Visual perception	Frame of reference	Visuomotor recorelation	Motor learning
Same as during adaptation ^a	+	+	+	+	+	+
Pointing at visual target without prisms	+		+	+	+	+
Pointing at visual target with unexposed hand		+	+	+		?
Verbal judgment of location of visual target		?	+	+		
Pointing at auditory target ^b	+			+		+
Verbal judgment of location of auditory target ^b				+		
Pointing straight ahead ^b	+					+
Pointing at visual target with different arm movements	+	+	+	+	?	
Pointing at visual targets in different locations	?	+	+	+	?	
Judgment of distance between hands ^b	+					
Judgment of location of passively moved adapted arm relative to visual target	+	+	+	+		
Pointing with adapted hand at unexposed hand ^b	+					+
Pointing with unexposed hand at adapted hand ^b	+					

Note.—The subject adapts by pointing with one arm, using a stereotyped arm movement, at a single target seen through prisms. A + indicates the prediction of an adaptive shift as large as that obtained with the task used during adaptation.

^a Except that (as with all the other tests) the subject cannot see his hand and receives no information about his accuracy.

^b While blindfolded.

subject would show no adaptive shift in pointing at targets).

4. *Visuomotor recorelation.* Visual perception does not change, but a given visual input is paired with a different motor output. Since only the visuomotor system used during adapta-

tion is altered, the unexposed arm and all nonvisual targets are unaffected.

5. *Motor-response learning.* The practiced arm acquires a new motor response to stimuli from a given spatial location regardless of their modality. There is a generalization de-

crement when the subject uses arm movements that differ from the practiced one.

Table 1 summarizes the predictions of the proprioceptive-change hypothesis and the other five conceptions. Several other, more sophisticated theories are discussed briefly at the end of this paper. Unfortunately, these theories often make equivocal predictions, or none whatever, about many of the tests listed in Table 1.

Experimental Findings. Harris (1963a, 1963b) carried out six of the tests listed in Table 1. The subjects, whose heads were held stationary by a bite board, adapted by pointing for 3 minutes at a visual target seen through prisms that displaced its image 11° to the right or left. Adaptation was found to produce sizable and significant adaptive shifts, which were virtually identical whether measured by pointing at visual targets, at auditory targets, or "straight ahead." The shift was no smaller when subjects pointed at targets several inches from the one they had practiced on, even though the arm movements used then differed from the practiced one. However, adaptation had little or no effect on pointing with the unexposed hand. Nor did it affect judgments of whether a given auditory target sounded straight ahead. (Hein & Held, 1960, had previously reported that, with a similar adaptation procedure, there was no change in judged location of visual targets.) Others have independently demonstrated the adaptive shift with auditory test targets and with pointing straight ahead (Pick, Hay, & Pabst, 1963) and the absence of any shift in pointing with the unexposed hand (H. B. Cohen, 1963; Hamilton, 1964a; Mikaelian, 1963; Scholl, 1926). Subsequent studies have also confirmed these three

findings (Goldstein²; Hay & Pick, in press; McLaughlin & Bower, 1965; McLaughlin & Rifkin, 1965).

On the basis of these results, five of the notions listed in Table 1 may be ruled out. The data can be accounted for only by the first interpretation: that adaptation consists of a change in the felt position of the adapted arm relative to the rest of the body.

Further Tests. The proprioceptive-change interpretation implies that a subject should make errors in judging how far his adapted hand is from other parts of his body—for example, his other hand. This prediction was tested by having subjects (who had adapted their right arms by pointing at a target seen through prisms) move their unexposed hands to specified subjective distances from their adapted hands while blindfolded (Harris, 1963a). After seeing their right hands shifted to the right by base-left prisms, subjects felt their hands to be farther apart, at a given physical distance, than when their hands were not adapted. Subjects who wore base-right prisms felt their hands to be closer together. These results demonstrate that there is in fact a change in the felt location of the adapted hand relative to the other hand.

During these tests, the subject was not allowed to make any active movements with his adapted arm. Thus, a simple motor-learning or conditioned-response theory of adaptation is inadequate: Although self-produced movements may be an essential precondition for adaptation, they are not a necessary part of the end product. The adaptive shift is evident whether the subject actively points at a target during the test or a luminous target is moved until he says it is right over his

² Donald Goldstein, personal communication, June 1964.

stationary hand (Hamilton & Hilliard³). A change in the position sense will indeed *affect* motor responses of the adapted arm, but the change does not itself *consist* of newly acquired motor habits. On the contrary, it is a *perceptual change* (in the felt position of the adapted arm).

The most direct evidence for the hypothesized change in position sense is obtained when the subject points with his unexposed arm (which points correctly at all other targets) at his stationary adapted arm. Harris' (1964) results fell just short of significance, but more recently Efstathiou and Held (1964) and Goldstein² have found large and significant shifts on this test, as well as on pointing with the adapted hand at the unexposed hand. Both findings were anticipated by Scholl (1926).

Related Findings. Several other recent studies fit in well with the proprioceptive-change hypothesis. Bossom and Hamilton (1963) and Hamilton (1964b) found that adaptation to displaced vision—in contrast to visual discrimination learning—shows complete interocular transfer and no intermanual transfer in split-brain monkeys; the adaptation is specific to the arm, not the eye. Nielsen (1963), Hay, Pick, and Ikeda (1965), Rock and Victor (1964), and Wertheimer and Arena (1959) have demonstrated that vision may immediately and completely dominate the position sense when the two disagree, a finding analogous to the smaller but longer lasting adaptive shifts discussed above.

All in all, it seems reasonable to conclude that, when a person watches one hand through prisms with little head movement, the adaptation is mainly a change in the felt position of that arm relative to the rest of his

body. Although their own data did not rule out all alternative hypotheses, other investigators have independently, at about the same time, reached similar conclusions: that such adaptation takes place in the "kinesthetic spatial system" (Hochberg, 1963; Pick et al., 1963) or, more specifically, in the adapted arm's position sense (Hamilton, 1964a, 1964b). A similar hypothesis was proposed earlier by Scholl (1926).

Head-Body Adaptation

Another way to adapt to displaced vision is simply to walk around while wearing prisms (Hay & Pick, in press; Held & Bossom, 1961; Kohler, 1951, 1964; Taylor, 1962). The results are quite different from those of arm adaptation. When presented with a visual target after the prisms are removed, the subject points incorrectly with *both* arms, even if he saw neither one through prisms (Bossom & Held, 1957), and says the target *looks* straight ahead of him when it is actually somewhat off to one side (Held & Bossom, 1961; Kohler, 1964).

Is this type of adaptation, then, completely unlike arm adaptation? Probably not. Just as the felt relationship between arm and body is altered by moving the arm while wearing prisms, so perhaps the felt relationship between head and body is altered by moving the head while wearing prisms.⁴ The three investigators who independently proposed this hypothesis

⁴ It is convenient to think of arm adaptation as a change in the felt position of the adapted arm, with that of the rest of the body remaining unchanged. However, the same phenomena would be observed if the perception of the arm stayed the same and that of all of the rest of the body changed. Strictly speaking, we can detect only a changed relationship between the two. This is even clearer in the case of head-body adaptation.

³ Charles R. Hamilton and S. A. Hilliard, personal communication, August 1964.

—Hamilton (1964a), Harris (1963a, 1963c), and Mittelstaedt (1964)—were unaware that Kohler (1951, p. 23) had already observed just such a phenomenon: A subject who wore prisms developed the "habit" of holding his head turned 6° – 9° to the right of his body midline but was "completely unaware" of the deviation.⁵ He felt that his head was pointing straight ahead.

It is the unawareness, not the turning, that is crucial. Contrary to Smith and Smith's claim (1962, pp. 92, 116–117), a "compensatory reaction" of turning the head cannot in itself counteract the prism-produced visual displacement: The perceived location of an object (relative to one's body) does not normally change when one turns one's head, because the new position of the head is taken into account. Perception of the object changes only if a subject misperceives the orientation of his head.

Whereas misperception of just one arm affects only tests that involve that arm, virtually any test will show the effects of misperceived head orienta-

⁵ "Als Prismenträger ist man ständig gezwungen, Auge und Kopf gegenüber der Greifrichtung etwas verdreht zu halten. Und in der Tat liess sich als Nachwirkung dieser aufgezwungenen 'Lebensgewohnheit' eine merkbare Verdrehung zwischen Kopf und Rumpf nachweisen, welche aber der Aufmerksamkeit der Vp. vollkommen entging. Sie meinte gerade und unverdreht zu stehen, während sie in Wirklichkeit den Kopf nach rechts gedreht hielt (6–9 Grad von der Körpermitte abweichend). Korrigierte man aber die Rechtslage des Kopfes, so entstand im Erleben der Vp. der Eindruck einer Linksverdrehung [cf. Kohler, 1964, p. 38]."

The Fiss and Gleitman translation (Kohler, 1964) differs, in many places, from the German papers on which it is based (Kohler, 1951, 1953). Though usually slight, the discrepancies are sometimes misleading at crucial points. Therefore, the present writer's translations are occasionally given instead.

tion. If a subject feels his head to be pointing straight ahead of his body when it is really somewhat turned, then when he sees an object directly in front of his nose he will incorrectly (if he is not wearing prisms) perceive that object to be straight ahead of his body. If he tries to point at it with either hand, he will point straight ahead of his body and thus point incorrectly. (Such misperception of head position would, of course, lead to improved accuracy of performance while the prisms are on.) Similar results will occur even if the test apparatus constrains the subject to hold his head straight relative to his body, as in Held and Bossom's (1961) procedure. When Kohler forced his subject to point his head straight ahead, the subject felt that it was turned several degrees to the left (1951, pp. 23–24).

A change in the felt relationship between head and body necessarily entails a change in the perceived direction of visual targets relative to the body. But it would be inaccurate to describe such adaptation as solely a change in visual perception, since, for example, altered perception of head orientation would also result in altered auditory localization.

Intermanual Transfer. A number of investigators have found that if a subject watches one hand through prisms, with little head movement, the adaptation is completely or almost completely confined to the exposed hand. Helmholtz (1962a, p. 157), however, reported considerable adaptation of the unexposed hand as well. How can these findings be reconciled?

A plausible answer was suggested independently by Hamilton (1964b) and Harris (1963a, 1963c). They both noted (as did H. B. Cohen, 1963) that subjects whose heads were immobilized while they adapted showed

little intermanual transfer, whereas those who were free to move their heads, as Helmholtz was, exhibited considerable transfer to the unexposed hand. Hamilton and Harris concluded that moving the head while wearing prisms leads to a change in the felt position of the head relative to the body, which would make the subject mispoint with both hands even if he never saw them through prisms. If he did see one arm through prisms, he would show a larger aftereffect with that arm than with the unexposed arm (since, in addition to the error caused by misperceiving the orientation of his head, there would also be a misperception of the exposed arm's orientation), thus manifesting "partial intermanual transfer."

If this analysis is correct, the term intermanual transfer is, in this context, something of a misnomer. Transfer implies that the adaptive change in one arm (or relevant parts of the nervous system) somehow spreads to or induces a similar change in the other arm (or contralateral part of the nervous system). Although this possibility has not been definitely ruled out (as Hamilton, 1964b, noted), it is simpler to assume that the measured adaptation in the unexposed arm results wholly from head-body adaptation, which affects both arms equally, and that there is in addition some arm adaptation of the exposed arm.

Wooster's Experiment. The concept of altered position sense of the head removes much of the mystery from a phenomenon reported by Wooster (1923). Her subjects reached beneath an opaque board to point at targets that were visible, through prisms, above the board. Surprisingly, even subjects who never saw their hands through prisms and so had no "knowledge of error" gradually became more accurate. Wooster considered a num-

ber of possible reasons for the improvement, tested them in further experiments, and found that none fitted all of her data.

Although Wooster's subjects did not walk around while wearing prisms, they were free to move their heads and so could have undergone head-body adaptation. Had Wooster tested the unpracticed arm, she might have been even more surprised to find that it had improved just as much as the practiced one.

A change in the proprioceptively perceived relationship between head and body could also account for many of the findings reported by Wallach, Kravitz, and Lindauer (1963), by Bossom (1964), and by McLaughlin and Rifkin (1965), as well as for the aftereffects of the incidental vertical displacements produced by Stratton's (1897, p. 471) and Kohler's (1964, p. 32) inverting goggles.

Eye-Head Adaptation

All of the phenomena ascribed above to head-body adaptation (except for Kohler's direct observation of a change in felt head position) might equally well be due to a change in the registered relationship between the eyes and the head (Harris, 1963a)—a modification of the "judgment of the direction of the gaze," as Helmholtz (1962b, p. 246) put it. Indeed, Kohler (1964, p. 32) says that in his experiments alterations in "kinesthetic sensations" from the eyes were often encountered.

Unlike head-body adaptation, eye-head adaptation would not affect auditory localization. Thus, a subject who misperceived the orientation of his eyes should misperceive the location of a visual target relative to that of a sound from an unseen source (Harris, 1963a). Recent studies have demonstrated just such an auditory-

visual mismatch. After certain adaptation procedures, a subject errs in judging where on a luminous visual scale a sound is coming from (Hay & Pick, in press; Wallach & Bernheim⁶). He points in one direction at a light and in another at a sound which is actually in the same place (Hay & Pick, in press; McLaughlin & Bower, 1965).

These and other findings reported by Hay and Pick (in press), McLaughlin and Rifkin (1965), and McLaughlin and Bower (1965), as they acknowledge in their later papers, are all attributable to a change in registered eye position plus a change in the exposed arm's position sense, with the amounts of the two changes varying during the course of adaptation.⁷

Because incorrect registration of eye position would entail incorrect localization of all seen objects, one can say that eye-head adaptation alters visual perception. But this alteration is fundamentally different from purely visual modifications such as dark adaptation and localized figural aftereffects. It is more akin to altered position sense in the arm or head. Thus it seems inadvisable to make a sharp distinction between "proprioceptive adaptation" (of the arm) and "visual adaptation." Such a distinction might, for example, lead to a fruitless search for modifications in the pathways connecting the retina to the visual cortex.

Half Prisms. Prisms that cover only the upper half of the eye displace the upper half of the visual field relative to the lower half, making a straight vertical line look discontinuous. Kohler (1964) reported that

subjects who adapt to half prisms say that the line eventually looks straight and unbroken most of the time despite the discontinuous retinal image.

Although this adaptation sounds like a purely visual change—a change in perceived relationships *within* the visual field—Kohler's other observations indicate otherwise. When an adapted subject was asked to move his eyes straight up and down in the dark, Kohler says, the subject actually moved them in a jagged line, with a sideways jump approximately in the middle of the movement ("*einen seitlichen Sprung ungefähr in der Mitte der Bewegung*"—Kohler, 1951, p. 73; cf. 1964, p. 93). With more rapid eye movements, the path became diagonal. But the subject "always thought that his eyes moved vertically and without sudden deflections [1964, p. 94]." Apparently, the subject perceived a broken line as straight only because he felt that the jagged eye movement he made in scanning it was straight. What happened when the subject *fixated* the dividing line between the prism and nonprism areas? Kohler (1964, p. 83) says explicitly that, when fixated, vertical lines looked just as discontinuous after many days of adaptation as they had at first. Clearly, there was no change in the purely "pictorial" aspect of visual perception, but only in those perceptions of visual location that depend on the registration of positions and movements of the eyes. Note that the adaptation did not entail any change in scanning *behavior*: When scanning the discontinuous line, the subject made essentially the same eye movements after adapting as he had before. The only change was that a jagged eye movement was interpreted as straight. This is a perceptual change, not the acquisition of new motor responses.

⁶ Hans Wallach and Joseph Bernheim, personal communication, December 1963.

⁷ Part of the shift that Hay and Pick (in press) attributed to arm adaptation may actually be due to head-body adaptation; Hay and Pick's tests do not distinguish between the two.

Curvature. Straight vertical lines look curved when viewed through a sideways-displacing prism because the prism displaces the top and bottom of the visual field more than the middle. The curvature is a set of relative displacements of the same sort as produced by a half prism. So perhaps adaptation to curvature also involves altered registration of eye movements without any change in scanning behavior. After adapting, the subject may feel that his eyes are moving in a straight line when they are actually tracing out a curve.⁸ Perhaps the "unstable aftereffect" experienced by one of Kohler's subjects ("straight objects—for example, long and heavy steel pipes—curved and straightened out while the amazed subject was in the very act of looking at them"—1964, pp. 37-38) was due to alternate scanning and fixation.

An analogous case of curvature adaptation, resulting from altered kinesthetic perception of movements of the arm, was recently studied in collaboration with Judith R. Harris (Harris, 1964). Subjects moved one hand back and forth along a horizontal straight line while looking through prisms that made the line look curved upwards or downwards. Before this practice and again afterwards, they were asked to draw straight horizontal lines while blindfolded. The prediction was that if a subject ran his hand along a straight line that looked curved upward, for example, a straight horizontal arm movement would come to *feel* curved upward, so the subject would compensate and draw a *downward* curve in order to feel that his

⁸ This idea was developed in conversations with Julian Hochberg. Experiments by M. M. Cohen (1963) and Held and Rekosh (1963) suggest that there are at least two kinds of curve adaptation; the registered eye-movement explanation applies to only one kind.

arm was moving in a straight line. A significant shift, in the predicted direction, was found. Note that this shift cannot be due to an intermodal figural aftereffect, because the shift is in the wrong direction; it cannot be motor learning, because the arm movements made during practice were actually straight.

ADAPTATION TO INVERTED RETINAL IMAGES

Is a proprioceptive-change interpretation appropriate when subjects adapt to optical transformations more drastic than displacement? Stratton's (1896, 1897) reports on his adaptation to "reinversion" of the retinal image indicate that the answer is yes.

Stratton's Experiments

Proprioceptive Changes. Stratton's reports are indeed difficult to comprehend—at times they sound bizarre, at times, self-contradictory. But it is clear that Stratton experienced proprioceptive changes similar to those considered above, though far more extensive and less stable.

When he first looked through inverting lenses, Stratton (1896) says,

... the parts of my body were *felt* to lie where they would have appeared had the instrument been removed; they were *seen* to be in another position. But the older tactual ... localization was still the *real* localization [p. 614].

Soon, however,

... the limbs began actually to feel in the place where the new visual perception reported them to be. ... The seen images thus became *real things* just as in normal sight. I could at length *feel* my feet strike against the *seen* floor, although the floor was seen on the opposite side of the field of vision from that to which at the beginning of the experiment I had referred these tactual sensations. I could likewise at times feel that my arms lay between my head and this new position of the feet; shoulders and head, however, which under the circumstances could never be directly seen, kept the old

localization they had had in normal vision, in spite of the logical difficulty that the shape of the body and the localization of hands and feet just mentioned made such a localization of the shoulders absurd [p. 615].

Proprioceptive changes such as these account for the behavioral aspects of Stratton's adjustment to inverted vision.⁹ If the felt locations and movements of his hands and feet came to agree with their seen locations and movements, he would have no trouble reaching for or kicking things, whereas before adapting he had to move the limb in a direction that felt wrong. When the new proprioceptive perceptions became stable enough to persist even when the limb was out of sight, responses with that limb would be completely normal with no need for conscious deliberation.

These proprioceptive changes also explain Stratton's feeling that he had achieved a "reharmonization" of touch and sight: Whenever he touched an object with an adapted limb, he felt it to be where he saw it, because he felt the limb to be in a new location that agreed with its visual location.

Upright Vision. Stratton was not primarily interested in behavioral adjustments nor in proprioceptive or intersensory alterations. He wanted to find out whether the usual (inverted) orientation of the retinal image is necessary for seeing things as upright. If so, Stratton (1896) said, "it is certainly difficult to understand how the scene as a whole could even temporarily have appeared upright when the retinal image was *not* inverted [p. 616]." Yet, he claimed, this was precisely what happened. After several

days of adaptation, the world seen through inverting lenses sometimes appeared to be "in normal position" (1896, p. 616), "right side up" (1897, pp. 358, 469), "rather upright than inverted" (1897, p. 354). Some psychologists have taken these statements as conclusive evidence of a change in Stratton's visual system. Others have maintained that Stratton's assertions mean nothing at all. Walls (1951), for example, insisted that Stratton's descriptions of the scene as "upright" were "entirely metaphorical" (p. 191) and that actually all that Stratton achieved was a harmony between current perceptions and inverted eidetic imagery of objects outside the field of view (p. 200).

Stratton himself, on the other hand, thought it was quite natural for things to come to look upright again since he believed that "harmony between touch and sight, . . . in the final analysis, is the real meaning of upright vision [1897, p. 475]." But although "harmony between touch and sight" might indeed make the perceived orientation of the body and of the seen world agree, both body and world might still be perceived as *inverted* rather than upright. Perceived uprightness must depend on some other factor. That factor, as many investigators have pointed out, is the sensations of pressure and tension in the feet, legs, and body produced by the pull of gravity. Recently, experiments on subjects with labyrinthine defects led Clark and Graybiel (1963) to suggest that such pressure cues, rather than labyrinthine cues, may in fact be the major determinants of the perceived direction of gravity. Under zero-gravity conditions, for instance, subjects perceive the direction of the surface that their feet are touching as downward (Simons, 1959).

When Stratton first put on invert-

⁹ Since Stratton's lens system rotated the retinal image through 180°, he actually adapted both to inversion and to reversal. Kohler's (1951, 1964) experiments, using mirrors that inverted the retinal image without reversing it, generally support Stratton's observations.

ing lenses, he felt gravity pulling *away* from the seen location of the floor; "the general feeling was that the seen room was upside down; the body of the observer . . . was felt as standard and as having an upright position [1897, p. 348]." But gradually he began to feel that his feet, then his legs and arms, then most of his body were all in "the place where the new visual perception reported them to be [1896, p. 615]." The new proprioceptive localization was not stable—sometimes he even seemed to feel his limbs in both the normal and the new locations at once (1897, pp. 345–346, 465)—but when his legs and body were clearly felt to be in the new place, so, of necessity, were the gravitational pulls. Because the direction of the pull of gravity is, by definition, *down*, objects seen to lie in the same direction from the head as the legs were felt to be were perceived as down. So the floor looked "down," making the room look "right side up."

Since Stratton's head and shoulders "kept the old localization they had had in normal vision," he should then have felt that his legs and body were not on the same side of his eyes as his chin and shoulders. In other words, his head should have felt inverted! This is evidently just what happened:

Outer objects . . . frequently seemed to be in normal position, and whatever there was of abnormality seemed to lie in myself, as if head and shoulders were inverted and I were viewing objects from that position, as boys sometimes do from between their legs [1896, p. 616].¹⁰

Stratton's simile conveys exactly the sense in which things seen through inverting lenses looked upright. They did *not* look the same as they did before the goggles were put on. Rather,

¹⁰ "At other times," Stratton (1896) noted, "the inversion seemed confined to the face or eyes alone [p. 616]."

they looked upright relative to the felt direction of gravity, the way things look when seen from between the legs. If you set a book upright on the floor and look at it from between your legs, you will see the bottom of the book as "down" and the top as "up." You will see the pointed part of a capital A above the open part. In this sense, everything will look upright. And yet, you will have trouble reading the book—the letters will look rather like upside-down print.

Kohler's Experiments

Kohler's accounts (1951, 1964) of adaptation to inversion help clarify the role of gravitational pulls on proprioceptively adapted body parts. When a partly adapted subject, who still saw the world as inverted, took hold of a string with a weight attached to the other end, he suddenly *saw* the weight as hanging from the string instead of floating upward like a balloon. The explanation may be that the hands and arms are often the first parts of the body to adapt (Taylor, 1962). So, when the weight pulled on the subject's arm and attracted his attention to it, he felt it pulling toward where he saw the floor and therefore perceived that direction as "down."

Several writers (e.g., Klein, 1960, p. 103) have assumed that gravitational cues provide a direct access to reality—a veridical standard to which visual perception, when shown the error of its ways, conforms. According to the present interpretation, however, gravitational cues will make the inverted scene look upright only if they are felt by some proprioceptively adapted body part. Prominent gravitational cues in an *unadapted* area (produced, for instance, by a weight hanging from the subject's chin) might

make the scene look even more clearly inverted.

If we assume that the adaptive change is in the felt direction of gravity, not in the visual system, we can make sense of Kohler's (1964) report that one subject, who had been wearing inverting goggles for several days, said that "two adjacent heads, one upright, the other inverted, were *both* perceived as upright [p. 32]." Apparently, one head (the physically erect one) looked upright in that its chin was seen below its forehead; the other one (normally oriented on the retina) looked more natural, more recognizable as a normal face.

Illusory Movements of the Visual Field

Ordinarily, when you move your head downward, objects enter your field of view from below and travel to the top. You perceive the external world as stationary. If you move your head downward while wearing inverting goggles, though, objects enter the visual field from the top and travel downward. As a result, the world appears to be moving rapidly downward. (With reversing goggles, sideways movements produce a similar illusion.) After a few days the illusory swinging diminishes, until eventually the world appears stationary during head movements (Kohler, 1964; Stratton, 1897; Taylor, 1962).

This sort of adaptation may also be more closely related to proprioceptive arm adaptation than to purely intra-visual phenomena. Stratton (1897, p. 358) noted that he saw the world as stationary only when he misperceived the direction in which *he* was moving: "Movements of the head or of the body . . . seemed to be toward that side on which objects entered the visual field, and not toward the opposite side," as they had felt when he first put on inverting lenses. This sounds

like a kinesthetic change, which would stabilize visual perception without any change in the neural mechanism that normally takes head movements into account to yield a stationary visual world; only the felt movement of the head, the input to this mechanism, changes.

ADAPTATION TO REVERSED RETINAL IMAGES

Kohler (1953, 1964) has described in detail how subjects who wear right-angle prisms, which reverse their retinal images right for left, eventually achieve normal behavior and what he calls "correct seeing" (1964, p. 140). At first reading, his account is as bewildering as Stratton's.

"Piecemeal" Adaptation

When a person puts on reversing prisms, Kohler says, he initially reaches in the wrong direction for things, makes wrong turns, and sees all writing as mirror writing. In attempting to cope with reversed vision, the subject tries out various tactics, such as deliberately heading left when his goal appears to be on his right. As the subject adapts behaviorally, during the course of several weeks, he becomes able to walk, reach, and turn correctly without resorting to such "tricks." Concurrently, Kohler claims, his visual perception changes in a peculiar piecemeal fashion: Some parts of the visual field are perceived correctly while other parts remain reversed. For example, after 18 days:

Inscriptions on buildings, or advertisements, were still seen in mirror writing, but the objects containing them were seen in the correct location. Vehicles driving on the "right" . . . carried license numbers in mirror writing . . . the subject is capable of localizing both sides of, say, a "3" correctly (open to the left, the curves to the right) and still see it mirrorwise [1964, p. 155]!

At this stage, even though the subject's spontaneous behavior is usually correct, he often becomes confused and makes "errors" when asked to attend to his "immediate visual experience."

After many weeks, Kohler says, the subject's behavior and vision are both reoriented. He achieves "almost completely correct impressions, even where letters and numbers were involved [1964, p. 160]."

When one thinks of adaptation as a change in visual perception, these observations are incomprehensible. How can vision ever be partly right way round, partly reversed?

Determinants of Judgments

In attempting to make sense of Kohler's puzzling observations, it is helpful to bear in mind four different determinants of what a subject says when the experimenter tries to find out whether he sees things right way round. The first two determinants are distinguishable aspects of perception, whereas the other two are essentially irrelevant to spatial perception. Doubtless, people differ in which of these factors enter into their judgments in a given situation, and a given person may judge differently at different times. But it is often possible to find out operationally which factors the subject's report is based on and to design experiments that avoid the ambiguities of previous reports.

1. *Directional perception.* When asked "Does that object appear to be on your right or on your left?" many people probably make a directional judgment relative to their dominant hand. If the object is seen to lie on the same side of the body as the right hand, the subject says: "It's on my right" or "on my right-hand side." The same kind of judgment can be elicited whether or not the words right

and left are used, whether or not the subject has a dominant hand, and whether or not he refers the judgment to his hand: The experimenter can simply touch any spot on the subject's body and ask whether an object appears to be on the same side of his body as the touched spot.

Such a judgment is based on one aspect of spatial perception—perception of the location of an object relative to some part of the body. This is the sort of perception that usually guides motor behavior such as reaching for an object or walking toward it (cf. the concept of "manipulable regions," Kohler, 1964, p. 163).

2. *Pictorial perception.* Most of the debate on adaptation to distorted vision has concerned this aspect of visual perception, though it has not been clearly differentiated from other determinants of subjects' judgments. Pictorial perception consists of "looking at" the "picture" received by the visual system (cf. Gibson's, 1950, concept of "the visual field"). It is most obvious in successive comparisons: For example, we can ask a subject whether an arrow he is looking at is pointing the same way as the locomotive in a painting he saw before the experiment began.

The perception of "clockwise" or "counterclockwise" motion and "east" or "west" on a map are probably pictorial perceptions for most people, based on purely visual memories. Thus, to test for changes in pictorial perception, we can keep a subject from seeing any clocks or maps during the experiment and then ask him whether something appears to be moving clockwise or counterclockwise, or whether it is on the same side as the 9 on a clock or the east coast on a map. As long as the subject has a visual image of a clock or a map, he need not even think about the labels "right" and

"left," nor about any part of his body, when making his judgment. (Occasionally a subject may make a directional judgment when asked to make a pictorial one—for example, if he remembers that the locomotive in the painting was "going toward my right"—that is, toward his right hand. But such exceptions have no theoretical importance once recognized for what they are.) Even when he uses the terms right and left, which are often characteristic of directional judgments, the subject may be making a pictorial judgment: Some people habitually think of a certain part of the visual field as "the lower left corner" without referring at all to any part of the body.

3. *Familiarity.* A subject often describes his first perceptions through reversing goggles as "strange," "unusual," "unfamiliar," "new," or "mirror imaged." Later he describes them as "normal," "natural," "all right," "usual," "familiar," "right way round" (see Kohler, 1964, p. 142). Such descriptions are based almost entirely on past experience with particular stimuli or classes of stimuli and can be changed through repeated visual observation, even without distorting spectacles. For example, a person (perhaps an apprentice typesetter) who practices reading mirror writing may soon say that it is beginning to look "natural" or "all right." But neither his directional nor his pictorial perception has changed. If asked to judge the location of part of a letter (relative to part of the body or to visual memories), a person gives the same answer whether the letter looks "familiar" or "strange."

4. *Labels.* It is risky to let a subject use the words right and left. First, the same word may be used to label two quite different sorts of perception, pictorial and directional.

Second, a subject wearing reversing prisms could decide to start calling everything right that he formerly called left, even if he had not adapted at all. And third, he may be inconsistent or hesitant about which word to use even when his perception is completely determinate, stable, and clearcut; as Kohler (1964) put it, "there are people who always have trouble when asked to tell quickly where right or left is, but who never have difficulty in reaching for seen objects [p. 153]." Labels like right and left do not affect perception; it is irrelevant that the subject has learned to call a certain direction left and another direction right.¹¹

Proprioceptive Changes

With these distinctions in mind, it is possible to reexamine Kohler's observations and conclude that adaptation to reversed vision can be ascribed to a radical change in the felt location of the arms, legs, and body relative to the head and eyes, without any change in pictorial visual perception.

Kohler (1953, p. 110; cf. 1964, p. 153), in fact, did observe some proprioceptive and kinesthetic changes during the course of adaptation to reversal. After several days of wearing the goggles, he reported, there was:

... a weakening of the right-left orientation of the body image, which becomes uncertain, especially in connection with movements that have been deliberately practiced in reverse. The subject may even turn left, with full

¹¹ Terms like "upright" and "inverted" are even more ambiguous. Saying that an object looks upright may mean that it appears to be in its usual position, that it looks the same as it did before inverting goggles were put on, that it looks the same as it feels, that its top is perceived to be pointing away from the direction of gravitational pulls, that its bottom appears to be near where one's feet are, or that it is oriented appropriately to the rest of the visual scene.

confidence, when he does a "right face" with his eyes closed. When he moves his head and hands, the kinesthetic position- and movement-sensations are completely in accord with the (reversed) visual field. Yet ultimately this leads to a "dead end" (two errors that cancel each other!).¹²

In a footnote, Kohler added:

... by touch, doors, for example, seem to open in a reversed direction (as compared with earlier), as if they had been turned around in the meantime. However, the pre-experimental "right-left" of the body image remains unchanged in the shoulder and upper-arm region, and from there it undertakes its new conquest. When the attention is concentrated on this region, there is almost never any error.¹³

Clearly, Kohler regarded these kinesthetic and proprioceptive changes as temporary aberrations of no theoretical importance, leading only to a "dead end." Since proprioception and vision were *both* reversed, Kohler thought that both must proceed to a further, "correct" stage before adaptation could be complete. Stratton's reports,

¹² "Aber auch umgekehrt schwächt sich das Rechts-Links der Körperfühlsphäre und wird unsicher, besonders im Zusammenhang mit jenen Bewegungen, die man absichtlich verkehrt eingeübt hat. Die Vp. kann mit voller Evidenz sogar bei geschlossenen Augen 'rechts un' machen und dreht sich dabei in Wirklichkeit nach links. Sie macht Kopfwendungen und Handbewegungen, deren kinästhetische Lage- und Bewegungsempfindung ganz mit der (verdrehen) visuellen Welt übereinstimmt. Was dabei letzten Endes herauskommt, führt aber in eine 'Sackgasse' (zwei Fehler, die sich gegenseitig aufheben!)."

¹³ "Das führt so weit, dass sogar im Tasten z. B. Türen (gegenüber früher) verkehrt aufzugehen scheinen, als wären sie inzwischen versetzt worden. Worauf sich das vorexperimentelle 'Rechts-Links' im Körpergefühl aber versteift und von wo es dann seinen neuen Vorstoß unternimmt, ist die Schulter- und Oberarmpartie. Wenn man darauf die Aufmerksamkeit konzentriert, gibt es kaum jemals eine Verwechslung."

(This footnote refers to the sentence that ends with "übereinstimmt.")

however, suggest that proprioceptive and kinesthetic changes, far from being temporary and trivial, become more and more extensive as adaptation progresses and are directly responsible for the "correct" perceptual judgments that ultimately emerge.

In order to clarify this interpretation of adaptation to reversal, let us consider a hypothetical experiment in which we test the subject's perception by having him look at a blackboard bearing an L on his left and an R on his right (Figure 1A). Immediately after putting on reversing spectacles, the subject says that he feels that his right hand (the one he writes with) and the right side of his body are near the same end of the blackboard—namely (since he is looking at it through reversing prisms), the end with a backwards L on it (Figure 1B). But when he holds up his writing hand and looks at it, he sees it nearer to the backwards R.

Now he starts adapting. If the proprioceptive-change hypothesis is correct, there is a change in the felt locations of his hands relative to his body. That is, his writing hand not only *looks* as if it is nearer to the backwards R, it now *feels* nearer as well (Figure 1C). Thus the subject feels his right hand to be near the (physically) left side of his body.

Suppose we ask the subject to turn right. He most likely assumes we mean toward the hand he writes with. Accordingly, he turns toward where he feels that hand to be, and the experimenter writes down that the subject turned left when told to turn right. The error is not due to uncertainty or "weakening of the right-left orientation of the body image"; if, instead of asking the subject to turn right, we touched his right hand and asked him to turn toward it, he would make the same error. If we

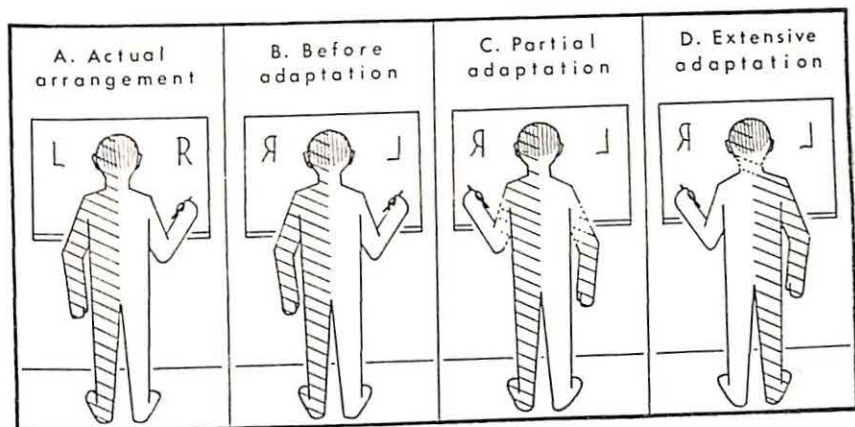


FIG. 1. A subject's perceptions during the course of adaptation to reversed vision, according to the proprioceptive-change hypothesis. (In all cases perception of the letters is visual; perception of the subject's head and subbody is proprioceptive. A: The actual physical arrangement. B: The subject's perceptions when he first puts on reversing goggles. C: The subject's perceptions at an intermediate stage of adaptation, with only his arms adapted. D: The subject's perceptions at an advanced stage of adaptation.)

touched his (unadapted) right shoulder, however, or if the subject focused his attention there, he would turn correctly to the right.

When we now ask the subject which end of the blackboard appears to be on his right, he may fall into the state of indecision that is so characteristic of Kohler's partly adapted subjects. His right hand feels closer to the letter R, whereas his right shoulder feels closer to the L. Depending on which region of his body he concentrates on, he can consider either the L or the R as being on his right. Thus he may switch his judgment back and forth from "right" to "left" without the slightest change in visual perception. Moreover, Stratton's reports indicate that the partly adapted subject sometimes feels his hand to be in two different places at once, with fluctuations in which of the localizations seems most "real." So even if the subject keeps judging relative to his hand, his judgments may waver. And, of course, he may switch back and forth from directional to pictorial judgments.

With further adaptation, according to the proprioceptive-change hypothesis, the felt locations of the subject's legs, torso, and perhaps even his shoulders and most of his head, change. He again feels, as he felt before the experiment started, that the (physically) right half of his body is near the hand he writes with. Now he can reach accurately for objects seen through the reversing goggles, turn correctly, and correctly judge the directions of objects relative to his body.

Pictorial Perception

But still his pictorial visual perception remains unchanged. The letters on the blackboard and the license plates on cars look just the same as they did when he first put on reversing goggles. The only difference is that he now feels the right side of his body to be on the same side as the curve of the backwards R that he sees (Figure 1D), so he says the curve is "on my right."

Why, then, does writing eventually come to look "normal" through reversing goggles? Because, with practice,

mirror writing becomes familiar and easy to read, whether the letters are actually printed in reverse or simply look reversed because one is wearing special goggles. Indeed, Kohler's (1964, p. 160) subject reported that "the first words to rectify themselves were the common ones," which were seen most often through the reversing goggles. But it is a mistake to conclude that pictorial perception reversed and "mirrorwise seeing" ("*spiegelbildliche Sehen*"—Kohler, 1953, p. 113) became established. We would not say that about someone who learned to read mirror writing without wearing reversing spectacles.

Given this interpretation of adaptation, it is not surprising to hear that one stubborn subject (Taylor, 1962, p. 180) "achieved satisfactory behavioral adjustment" but "denied that he ever perceived the world the right way round through his spectacles," even after 71 days of wearing them. When questioned closely, this subject said that all of his "incorrect" judgments were made by deciding whether his right temple or his left temple was closer to the part of the visual field he was judging. Stratton's reports suggest that even after the felt location of the rest of the body has changed, the area around the eyes does not, so judgments made relative to the temples would remain unchanged. Or, the subject may have been trying to tell the experimenter that he was making a pictorial judgment, based on an unaltered visual memory of right and left. In either case, this subject's perceptions—visual and proprioceptive—were probably just the same as those of subjects who, according to Taylor, managed to "perceive the world the right way round."

Aftereffects

The "peculiar experiences" (Kohler, 1964, p. 158) that the adapted sub-

ject encounters when he takes off the experimental spectacles are just what one would expect if the subject has undergone proprioceptive changes, has become accustomed to the reversed appearance of particular stimuli, but has experienced no change in pictorial perception.

When Kohler (1964, p. 158) removed his reversing goggles, after weeks of adaptation, and looked at a picture which he had seen before but not during the experiment, the picture immediately looked familiar. The person in it appeared (pictorially) to be running, as before, from left to right. Nevertheless, the person was seen as running toward the *left* edge of the page; that is (as Kohler makes clear), toward where Kohler felt his (adapted) left shoulder to be.

Kohler (1964, p. 160) does report that another subject, the one who "achieved almost completely correct impressions" while wearing reversing goggles, *saw* the whole room mirrorwise when the spectacles were removed. But the evidence for this statement is that the subject read p's as q's, b's as d's, and 10:30 on a clock as 1:30—which is just what would happen, without any perceptual change, if one read nothing but mirror writing and saw nothing but backwards clocks for 37 days.

LIMITATIONS OF THE PROPRIOCEPTIVE-CHANGE HYPOTHESIS

Although changes in the position sense may underlie most of the phenomena of adaptation to optical distortions, there are some kinds of adaptation that cannot be so interpreted. For example, adaptation to the chromatic dispersion produced by prisms ("color fringes"—Kohler, 1964) seems to depend on changes in "contour detectors" within the visual system (Hay, Pick, & Rosser, 1963; McCol-

lough, 1965a). Adaptation to bicolored spectacles (Kohler, 1964) has been shown to depend on retinal color adaptation and simultaneous contrast (McCollough, 1965b). There appear to be several forms of adaptation to tilting, curvature, and other optical distortions (see, e.g., M. M. Cohen, 1963; Held & Rekosh, 1963; Kohler, 1964; Mikaelian & Held, 1964; Morant & Harris, 1965; Ohwaki, 1961). Some of these appear to be purely visual changes; others may be based on changed registration of head or eye positions and movements.

Many adaptation situations probably include some motor-skill learning. This component may sometimes be even larger than that due to proprioceptive changes (as, perhaps, in Smith & Smith's, 1962, studies) or it may be much smaller but still considerable. For example, Hall (1964) and Harris (unpublished data) both found that when the arm movement used during adaptation differed grossly from that used in the tests, there was some decrement in the measured adaptive shift. (Hall's data have been published by Freedman, Hall, & Rekosh, 1965.)

In some cases, a proprioceptive-change interpretation requires additional assumptions similar to those made by other theories. For instance, the increased variability of responses that follows watching one's hand through a variable prism whose amount of displacement keeps changing (Cohen & Held, 1960; summarized by Held & Freedman, 1963) could be due to increased uncertainty about arm position. Many of the "conditional aftereffects," which others ascribe to the conditioning of visual perceptions to nonvisual cues (Kohler, 1964; Taylor, 1962), can be attributed to the conditioning of altered position sense to these same cues.

Recently, several experimental findings have been cited as directly ruling

out a proprioceptive change in adaptation to displacement. Efstathiou and Held (1964) reported that adapting one hand to displacing prisms had no effect on blindfolded subjects' reaching for the remembered locations of targets they had previously felt. They also found that the measured adaptive shift was smaller when the unexposed hand served as test target than when the target was a visually perceived object. Bauer and Efstathiou (1965) found an adaptive shift in pointing "straight ahead" only if subjects were first tested on pointing at visual targets; if tested on straight ahead first, the shift was sizable, but in the wrong direction. And H. B. Cohen (1963) reported that adaptation with the target in the retinal periphery does not transfer completely to test targets on the fovea. It is difficult to assess these findings since each is contradicted, directly or indirectly, by other findings (e.g., Goldstein¹⁴; Hamilton & Hillyard¹⁵; Harris, 1963a). The reasons for the empirical disagreements have not been satisfactorily worked out.

In the reports of Stratton (1897), Kohler (1964), and Taylor (1962), several passages seem to describe perceptual changes that cannot be attributed to altered position sense. Further work is needed to determine whether these statements are based on confusions about the various determinants of subjects' verbal reports or result from some complicated alterations in position sense or do in fact represent other sorts of adaptive change.

OTHER THEORIES OF ADAPTATION

Stratton

After reading Stratton's striking descriptions of the proprioceptive changes

¹⁴ Donald Goldstein, personal communication, June 1964, May 1965.

¹⁵ Charles R. Hamilton and S. A. Hillyard, personal communication, July 1964, August 1964.

he underwent, one is surprised to find him saying in his theoretical discussion that "the tactual perceptions, as such, never changed their place," and "the restoration of harmony between the perceptions of sight and those of touch was in no wise a process of changing the absolute position of tactual objects so as to make it identical with the place of the visual objects [1897, p. 476]." He seems to be ignoring his own introspections when he claims that there is neither a change in proprioceptive localization nor a change in visual localization, but only a change in the relationship between the two. This noncommittal idea has proven quite appealing to many present-day psychologists.

Clearly a change either in vision or in the position sense would result in a changed relationship between the two. But saying that only the relationship between the two modalities changes is ignoring information about the changes *within* one (and only one) of the modalities.

Stratton's own reports make it clear that the first step in his adapting to inverted vision was to feel that his feet were in a new location relative to the rest of his body—a change within the position sense. Gradually, the felt locations of more and more of his body swung into line with that of his feet, that is, into line with the inverted visual scene. True, the final result was a new relationship between the position sense and visual perception, but this new relationship was brought about entirely by changes in the position sense, with no changes in vision.

In 1897, Stratton (pp. 472-475) theorized that adaptation is the attachment of new visual imagery to tactual sensations and, concurrently, the attachment of new tactual imagery to visual sensations. If we ignore the

second half of this formulation and just postulate that adaptation consists of associating new visual imagery of parts of the body with proprioceptive stimuli from those parts, we can deal with much of the relevant data, provided we make one additional assumption: that the felt position of a limb is not directly connected with proprioceptive stimuli, but is a byproduct of where the limb is mentally pictured. The visual imagery notion, then, would make the same predictions as the proprioceptive-change hypothesis, but requires an extra step—a step that some subjects' introspections deny.¹⁶

Taylor

Taylor (1962, 1964) attempts to explain all perception, from depth perception to color vision, with a single hypothetico-deductive theory. According to this theory, visual perception of an object is determined by the "activation" of stimulus-response engrams—neural traces of the responses (especially walking, reaching, and verbal responses) that have been conditioned to similar stimuli. Taylor apparently believes that adaptation to displacement, inversion, or reversal of the retinal images leads to changes in directional perception and perhaps to pictorial changes as well (1964, p. 73). However, he thinks that these perceptual changes, though genuine, are largely

¹⁶ In a later paper, Stratton (1899) did state, contradicting his earlier theoretical views, that "the place in which any part of the body is persistently seen influences the localisation of the dermal and kindred sensations arising in that part. If one were to see his feet, for instance, in some direction different from their present visual position, he would in the end refer thither their kin-aesthetic impressions also [p. 463]." But only a few subsequent writers (notably Walls, 1951; Smith & Smith, 1962) have paid much attention to Stratton's later interpretation or even to his original detailed descriptions of proprioceptive changes.

the result of changes in verbal labeling (1962, pp. 179–181, 185). He expects adaptation to progress in piecemeal fashion with visual perception becoming more and more veridical as the subject acquires a larger number of appropriate responses (1962, pp. 188, 197, 207).

Beyond this, it is difficult to derive unequivocal predictions. For example, Taylor's theory could predict either way about transfer to most of the tasks in Table 1, depending on what assumptions are made about steepness of generalization gradients, breadth of "equivalence classes," degree of "interpenetration" of sensory modalities, and relative importance of motor behavior and implicit verbal responses.

Kohler

Kohler's (1964) studies of adaptation to a wide variety of visual distortions have provided the inspiration, directly or indirectly, for much of the research in this field. He has concentrated on setting down his observations rather than on providing a detailed theory. It is clear that he agrees with Taylor that adaptation involves changes in directional perception, based on the acquisition of new behavioral responses to transformed retinal images (see, e.g., pp. 163–164). With prolonged exposure to reversing spectacles, Kohler says, there are eventually pictorial changes, with more and more stimuli "seen correctly" (pp. 140, 163–164). Unlike Taylor, however, Kohler thinks that verbal labeling is of no great significance in adaptation.

Although Kohler did mention (in a footnote) that "alterations in kinesthetic sensitivity" may be "of crucial importance" (p. 32) in adapting to displacement, he did not make much use of these alterations in explaining other aspects of adaptation. In fact, in his dis-

cussions of reversed vision, he regards such alterations as transitory—normal proprioception and kinesthesia are soon reinstated, and form the basis for the "correct" visual perception that ultimately emerges. In his theoretical discussions, Kohler did not attempt to explain the simpler phenomena of adaptation to displacement, inversion, and reversal of retinal images, but rather dealt with the complex "situational aftereffects."

Held

In an extensive series of carefully controlled experiments, Held and his co-workers have demonstrated the importance of active movement and movement-produced visual feedback ("reference") in producing adaptation. These experiments set the pattern for most of the recent work in the area: brief adaptation periods with quantitative before-after measurements.

Held has been primarily concerned with the necessary preconditions for adaptation; he has said little about the nature of the adaptive change (see, e.g., Held, 1961). It is not clear whether Held believes that adaptation involves any perceptual changes, visual or proprioceptive. For instance, Held and Freedman (1963) say that adaptation "represents a change in state of the relevant sensorimotor control system, such that [after complete adaptation] the input-output or stimulus-response relation becomes identical to that which existed prior to rearrangement [p. 457]." Recently Efstathiou and Held (1964) proposed a tentative theory of arm adaptation to displacement, according to which "the change responsible for the shifts occurs in a representation, within the nervous system, of the spatial relation between the exposed arm and directions that are defined independently of that arm." Further elaboration of this model is necessary to

determine how it differs from the proprioceptive-change interpretation.

Smith and Smith

Smith and Smith (1962) claim that adaptation consists of acquiring highly specific perceptual-motor skills. They seem to deny that there is any general reorientation of perception, whether directional or pictorial (1962, pp. 83, 311). Their research and theory supplement the proprioceptive-change interpretation by dealing with situations in which proprioceptive changes probably are minimized because there is a large spatial separation between the felt location of the hand and its televised visual feedback. The acquisition of highly specific perceptual-motor skills is facilitated by Smith and Smith's tasks, which permit continuous visual feedback and stress speed of execution. Indeed, their usual measure (speed of performance) is sensitive only to the development of highly practiced motor skills. However, it is possible that adaptation of Smith and Smith's three movement systems—locomotion, transport, and manipulation—may in part represent, respectively, proprioceptive head or eye adaptation, arm adaptation, and acquisition of manipulatory skills.

Werner and Wapner

Werner and Wapner (1955) have discussed some of Kohler's findings in terms of their organismic sensory-tonic field theory, attributing adaptation to changes in the subject's "organismic state (sensory-tonic distribution)." Basically, though, their account simply restates Kohler's observations. Some of their statements would match the present author's if the words "felt position" of certain body parts were substituted for such abstract terms as "organismic state" or "equilibrical axis." But Werner and Wapner consider the

organismic state to be only one part of the process that determines body perception, not the perception itself. Moreover, the organismic state is assumed to include "not only postural, but emotive, motivational factors, etc. [p. 133]," which are clearly beyond the scope of the present formulation.

PRECONDITIONS AND MECHANISM FOR ADAPTATION

Visual Proprioceptive Discrepancy

Although the proprioceptive-change interpretation of adaptation does not specify any particular process or precondition for the change, it is tempting to assume that adaptation results from a discrepancy between proprioceptive and visual information. One effective way to produce such a discrepancy is to look at some part of one's body through distorting goggles, but it is not the only way. For instance, when a subject walks while wearing displacing prisms, his position sense may indicate that his head is turned to one side of the direction of movement, whereas the retinal flow pattern (Gibson, 1950; Held & Freedman, 1963) may indicate that the head is pointing right along the axis of movement.

On the other hand, there is no logical necessity that proprioceptive inputs be used at all by the mechanism that recalibrates the position sense. It could be, as Helmholtz (1962b) suggested, that adaptation is based on the changes in the retinal image that result from a given "effort of the will," or, in Held and Freedman's (1963) terminology, on motor corollary discharges and visual reafference contingent upon active movement. Or the proprioceptive change could be due to the laying down of engrams of conditioned responses, like those postulated by Taylor (1962).

Some proprioceptive changes may

arise from something like "sensory fatigue," without any direct participation by vision. Hein (1965) has found that after simply holding their heads turned to one side for 10 minutes, even with their eyes closed, subjects point incorrectly at visual targets. This "postural after-effect," as Hein called it, probably was due to a change in felt head orientation. Similarly, Kohler (1951, p. 18) reports that subjects who wear displacing prisms tend to hold their eyes in an abnormal position that eventually comes to feel normal once more. Such a "fatigue" effect could indeed produce some of the phenomena of adaptation to displacement, but it can underlie neither adaptation to inversion or reversal nor arm adaptation to displacement.

Active Movement

A number of experiments by Held and his colleagues have shown that active movements by the subject play an important role in adaptation to several optical distortions (Held & Freedman, 1963). Although some recent investigators have claimed to find extensive adaptation with passive exposure (e.g., Wallach et al., 1963; Weinstein, Sersen, Fisher, & Weisinger, 1964), active movement does seem greatly to facilitate adaptation.

Theories that consider the end product of adaptation to consist of new motor responses or new visuomotor correlations (e.g., Held, 1961; Smith & Smith, 1962; Taylor, 1962) also assume that active movement is a crucial precondition for adaptation. On the other hand, Hamilton (1964a) has listed several ways to account for the importance of active movement without postulating any motoric component in the end product. For example, the position sense during active movement may differ from (and be more precise than) that during passive movement.

Or motor discharges may act as a "catalyst" that permits a joint's position sense to change.

IMPLICATIONS FOR PERCEPTUAL DEVELOPMENT

Psychologists have traditionally looked to studies of adaptation to distorted vision for clues about the development of visual perception in the infant. The usual, empiricist assumption (outlined by Berkeley in 1709 in his *New Theory of Vision*; see Berkeley, 1910) is that visual space perception is "secondary": It is based on the spatial sensations given by touch, kinesthesia, and position sense. As Dewey (1898) put it: "Ultimately visual perception rests on tactual. . . . Spatial relations are not originally perceived by the eye, but are the result of the association of visual sensations with previous muscular and tactual experiences [p. 165]."

This belief in the primacy of touch is so ingrained that experimental results are sometimes flagrantly misinterpreted in order to support it. Carr (1925), for instance, concluded: "It is thus obvious that the Stratton experiment involves no reconstruction or alteration of tactual . . . space. It is the visual system that is disrupted and then reorganized so as to conform to touch . . . [p. 141]." Stratton's, Kohler's, and Held's findings have been cited over and over as evidence that visual space perception is flexible and therefore must have been acquired through tactile-proprioceptive and motor experience. The reinterpretation of these findings that has been presented here suggests the opposite conclusion. Vision seems to be largely inflexible, whereas the position sense is remarkably labile.

The implication, if one dare draw any, is that the Berkeleyan notion should be turned around. It seems more plausible to assume that proprio-

ceptive perception of parts of the body (and therefore of the locations of touched objects) develops with the help of innate visual perception rather than vice versa.¹⁷ A growing number of recent studies support the view that many aspects of visual perception are not influenced by experience and are largely innate (e.g., Bower, 1964; Fantz, 1965; Gibson & Walk, 1960; Hubel & Wiesel, 1963; Robinson, Brown, & Hayes, 1964). Furthermore, if the position sense were innate—if each spot on the skin were proprioceptively “preaddressed”—the local sign lodged in a baby’s fingertip might go on forever signaling that his arm is 10 inches long.

So, when a baby stares raptly at his outstretched hand, he is probably find-out where his hand is, not what his visual sensations mean. He is making use of an adaptive mechanism that keeps his position sense accurate despite extensive and uneven growth of his body. This mechanism enables us to use the precise, detailed information that vision provides, as a means of continually readjusting our vaguer and more variable position sense.

¹⁷ Clearly vision is not the *only* basis for acquiring and maintaining the position sense or blind people would have no idea where their arms and legs were. Vision may, however, provide the quickest and most exact recalibration.

REFERENCES

- BAUER, J., JR., & EFSTATHIOU, AGLAIA. Effects of adaptation to visual displacement on pointing “straight ahead.” Paper read at Eastern Psychological Association, Atlantic City, April 1965.
- BERKELEY, G. *An essay towards a new theory of vision*. New York: Dutton, 1910.
- BOSSOM, J. Mechanisms of prism adaptation in normal monkeys. *Psychonomic Science*, 1964, 1, 377-378.
- BOSSOM, J., & HAMILTON, C. R. Interocular transfer of prism-altered coordinations in split-brain monkeys. *Journal of Comparative and Physiological Psychology*, 1963, 56, 769-774.
- BOSSOM, J., & HELD, R. Shifts in egocentric localization following prolonged displacement of the retinal image. *American Psychologist*, 1957, 12, 454. (Abstract)
- BOWER, T. G. R. Discrimination of depth in premotor infants. *Psychonomic Science*, 1964, 1, 368.
- BRINDLEY, G. S., & MERTON, P. A. The absence of position sense in the human eye. *Journal of Physiology*, 1960, 153, 127-130.
- CARR, H. A. *Psychology: A study of mental activity*. New York: Longmans, Green, 1925.
- CLARK, B., & GRAYBIEL, A. Perception of the postural vertical in normals and subjects with labyrinthine defects. *Journal of Experimental Psychology*, 1963, 65, 490-494.
- COHEN, H. B. Transfer and dissipation of aftereffects due to displacement of the visual field. *American Psychologist*, 1963, 18, 411. (Abstract)
- COHEN, M. M. Visual curvature and feedback factors in the production of prismatically induced curved line aftereffects. Paper read at Eastern Psychological Association, New York, April 1963.
- COHEN, M., & HELD, R. Degrading visual-motor coordination by exposure to disordered re-afferent stimulation. Paper read at Eastern Psychological Association, New York, April 1960.
- DEWEY, J. *Psychology*. (3rd ed.) New York: American, 1898.
- EFSTATHIOU, AGLAIA, & HELD, R. Cross-modal transfer of adaptation to eye-hand rearrangement. Paper read at Eastern Psychological Association, Philadelphia, April 1964.
- FANTZ, R. L. Ontogeny of perception. In A. M. Schrier, H. F. Harlow, & F. Stollnitz (Eds.), *Behavior of nonhuman primates*. New York: Academic Press, 1965. Pp. 365-403.
- FREEDMAN, S. J., HALL, SARAH B., & REKOSH, J. H. Effects on hand-eye coordination of two different arm motions during compensation for displaced vision. *Perceptual and Motor Skills*, 1965, 20, 1054-1056.
- GIBSON, ELEANOR J., & WALK, R. D. The “visual cliff.” *Scientific American*, 1960, 202(4), 64-71.
- GIBSON, J. J. *The perception of the visual world*. Boston: Houghton Mifflin, 1950.
- HALL, SARAH B. Transfer of adaptation within the arm. Unpublished senior honors thesis, Harvard University, 1964.

- HAMILTON, C. R. Intermanual transfer of adaptation to prisms. *American Journal of Psychology*, 1964, 77, 457-462. (a)
- HAMILTON, C. R. *Studies on adaptation to deflection of the visual field in split-brain monkeys and man*. (Doctoral dissertation, California Institute of Technology) Ann Arbor, Mich.: University Microfilms, 1964, No. 64-11,398. (b)
- HARRIS, C. S. *Adaptation to displaced vision: A proprioceptive change*. (Doctoral dissertation, Harvard University) Ann Arbor, Mich.: University Microfilms, 1963, No. 63-8162. (a)
- HARRIS, C. S. Adaptation to displaced vision: Visual, motor, or proprioceptive change? *Science*, 1963, 140, 812-813. (b)
- HARRIS, C. S. The nature of adaptation to displaced vision. Paper read at Eastern Psychological Association, New York, April 1963. (c)
- HARRIS, C. S. Proprioceptive changes underlying adaptation to visual distortions. *American Psychologist*, 1964, 19, 562. (Abstract)
- HAY, J. C., & PICK, H. L., JR. Visual and proprioceptive adaptation to optical displacement of the visual stimulus. *Journal of Experimental Psychology*, 1966, 71, in press.
- HAY, J. C., PICK, H. L., JR., & IKEDA, KAREN. Visual capture produced by prism spectacles. *Psychonomic Science*, 1965, 2, 215-216.
- HAY, J. C., PICK, H. L., JR., & ROSSER, EDWENNA. Adaptation to chromatic aberration by the human visual system. *Science*, 1963, 141, 167-169.
- HEIN, A. Postural after-effects and visual-motor adaptation to prisms. Paper read at Eastern Psychological Association, Atlantic City, April 1965.
- HEIN, A. V., & HELD, R. M. Transfer between visual-motor systems of adaptation to prismatic displacement of vision. Paper read at Eastern Psychological Association, New York, April 1960.
- HELD, R. Exposure-history as a factor in maintaining stability of perception and coordination. *Journal of Nervous and Mental Disease*, 1961, 132, 26-32.
- HELD, R., & BOSSOM, J. Neonatal deprivation and adult rearrangement: Complementary techniques for analyzing plastic sensory-motor coordinations. *Journal of Comparative and Physiological Psychology*, 1961, 54, 33-37.
- HELD, R., & FREEDMAN, S. J. Plasticity in human sensorimotor control. *Science*, 1963, 142, 455-462.
- HELD, R., & REKOSH, J. Motor-sensory feedback and the geometry of visual space. *Science*, 1963, 141, 722-723.
- HELMHOLTZ, H. VON. *Popular scientific lectures*. (Ed. by M. Kline) New York: Dover, 1962. (a)
- HELMHOLTZ, H. VON. *Treatise on physiological optics*. (Trans. & Ed. by J. P. C. Southall) Vol. 3. New York: Dover, 1962. (b)
- HOCHBERG, J. On the importance of movement-produced stimulation in prism-induced after-effects. *Perceptual and Motor Skills*, 1963, 16, 544.
- HUBEL, D. H., & WIESEL, T. N. Receptive fields of cells in striate cortex of very young, visually inexperienced kittens. *Journal of Neurophysiology*, 1963, 26, 994-1002.
- KLEIN, G. S. Cognitive control and motivation. In G. Lindzey (Ed.), *Assessment of human motives*. New York: Grove Press, 1960. Pp. 87-118.
- KOHLER, I. Über Aufbau und Wandlungen der Wahrnehmungswelt. *Österreichische Akademie der Wissenschaften, Sitzungsberichte, Philosophisch-historische Klasse*, 1951, 227, 1-118.
- KOHLER, I. Umgewöhnung im Wahrnehmungsbereich. *Die Pyramide*, 1953, 3, 92-96, 109-113, 132-133.
- KOHLER, I. The formation and transformation of the perceptual world. (Trans. by H. Fiss) *Psychological Issues*, 1964, 3(4).
- LOYD, ANDREE J., & CALDWELL, L. S. Accuracy of active and passive positioning of the leg on the basis of kinesthetic cues. *Journal of Comparative and Physiological Psychology*, 1965, 60, 102-106.
- LUDVIG, E. Possible role of proprioception in the extraocular muscles. *Archives of Ophthalmology*, 1952, 48, 436-441.
- MCCOLLOUGH, CELESTE. Color adaptation of edge-detectors in the human visual system. *Science*, 1965, 149, 1115-1116. (a)
- MCCOLLOUGH, CELESTE. The conditioning of color perception. *American Journal of Psychology*, 1965, 78, in press. (b)
- MCLAUGHLIN, S. C., & BOWER, J. L. Auditory localization and judgments of straight ahead during adaptation to prism. *Psychonomic Science*, 1965, 2, 283-284.
- MCLAUGHLIN, S. C., & RIFKIN, K. I. Change in straight ahead during adaptation to prism. *Psychonomic Science*, 1965, 2, 107-108.
- MIKAELIAN, H. Failure of bilateral transfer in modified eye-hand coordination. Paper

- read at Eastern Psychological Association, New York, April 1963.
- MIKAELIAN, H., & HELD, R. Two types of adaptation to an optically-rotated visual field. *American Journal of Psychology*, 1964, 77, 257-263.
- MITTELSTAEDT, H. The role of movement in the origin and maintenance of visual perception: Discussion. In *Proceedings of the Seventeenth International Congress of Psychology*. Amsterdam: North-Holland, 1964. P. 310. (Abstract)
- MORANT, R. B., & HARRIS, JUDITH R. Two different after-effects of exposure to visual tilts. *American Journal of Psychology*, 1965, 78, 218-226.
- MOUNTCASTLE, V. B., POGGIO, G. F., & WERNER, G. The relation of thalamic cell response to peripheral stimuli varied over an intensive continuum. *Journal of Neurophysiology*, 1963, 26, 807-834.
- NIELSEN, T. I. Volition: A new experimental approach. *Scandinavian Journal of Psychology*, 1963, 4, 225-230.
- OHWAKI, SONOKO. An investigation of figural adaptation: A study within the framework of sensory-tonic field-theory. *American Journal of Psychology*, 1961, 74, 3-16.
- PICK, H. L., JR., HAY, J. C., & PABST, JOAN. Kinesthetic adaptation to visual distortion. Paper read at Midwestern Psychological Association, Chicago, May 1963.
- ROBINSON, J. S., BROWN, L. T., & HAYES, W. H. Test of effects of past experience on perception. *Perceptual and Motor Skills*, 1964, 18, 953-956.
- ROCK, I., & VICTOR, J. Vision and touch: An experimentally created conflict between the two senses. *Science*, 1964, 143, 594-596.
- ROSE, J. E., & MOUNTCASTLE, V. B. Touch and kinesthesia. In J. Field (Ed.), *Handbook of physiology*. Section 1. *Neurophysiology*. Vol. 1. Washington, D. C.: American Physiological Society, 1960. Pp. 387-429.
- SCHOLL, K. Das räumliche Zusammenarbeiten von Auge und Hand. *Deutsch Zeitschrift für Nervenheilkunde*, 1926, 92, 280-303.
- SIMONS, J. C. Walking under zero-gravity conditions. United States Air Force, Wright Air Development Center, Technical Note No. 59-327, 1959. Cited by J. P. Loftus, Jr., & Lois R. Hammer. Weightlessness. In N. M. Burns, R. M. Chambers, & E. Hendler (Eds.), *Unusual environments and human behavior*. New York: Free Press, 1963. Pp. 353-377.
- SMITH, K. U., & SMITH, W. K. *Perception and motion*. Philadelphia: Saunders, 1962.
- STRATTON, G. M. Some preliminary experiments on vision without inversion of the retinal image. *Psychological Review*, 1896, 3, 611-617.
- STRATTON, G. M. Vision without inversion of the retinal image. *Psychological Review*, 1897, 4, 341-360, 463-481.
- STRATTON, G. M. The spatial harmony of touch and sight. *Mind*, 1899, 8, 492-505.
- TAUB, E., ELLMAN, S. J., & BERMAN, A. J. Conditioned grasp response in a deafferented primate limb. *American Psychologist*, 1964, 19, 510. (Abstract)
- TAYLOR, J. G. *The behavioral basis of perception*. New Haven: Yale Univ. Press, 1962.
- TAYLOR, J. G. What is consciousness? *British Journal of Statistical Psychology*, 1964, 17, 71-76.
- WALLACH, H., KRAVITZ, J. H., & LINDAUER, JUDITH. A passive condition for rapid adaptation to displaced visual direction. *American Journal of Psychology*, 1963, 76, 568-578.
- WALLS, G. L. The problem of visual direction. *American Journal of Optometry*, 1951, 28, 55-83, 115-146, 173-212.
- WEINSTEIN, S., SERSEN, E. A., FISHER, L., & WEISINGER, M. Is reafference necessary for visual adaptation? *Perceptual and Motor Skills*, 1964, 18, 641-648.
- WERNER, H., & WAPNER, S. The Innsbruck studies on distorted visual fields in relation to an organismic theory of perception. *Psychological Review*, 1955, 62, 130-138.
- WERTHEIMER, M., & ARENA, A. J. Effect of exposure time on adaptation to disarranged hand-eye coordination. *Perceptual and Motor Skills*, 1959, 9, 159-164.
- WOOSTER, MARGARET. Certain factors in the development of a new spatial coordination. *Psychological Monographs*, 1923, 32(4, Whole No. 146).

(Early publication received June 10, 1965)

STRUCTURE OF PHENOMENAL DOMAINS

JOHN L. RINN¹

Division of Counseling Psychology, University of California, Berkeley

The development of theoretical systems in the social sciences has proceeded without sufficient attention to structural similarities between different universes of content. This has interfered with the discovery of isomorphic relationships between theories. In this paper the central variables of teacher-counselor personality are assigned to 3 domains (Personal, Interpersonal, and Attitudinal) which are analyzed in terms of similar dimensional structures. Facet analysis is used to compare domain elements and to discover empirical patterns of variables. Finally, the interdomain correlations are used to construct a heuristic typology of counselor syndromes.

When Jimmy puts a mouse into teacher's desk, many things are likely to happen. If I describe this event from my point of view and you describe it from yours, the two descriptions will certainly be different and may very well appear to refer to two separate events. Among other explanations for this occurrence is the one that you and I are attending to different classes of phenomena, that we perceive and describe events selectively.

How then can we communicate with and understand each other? My proposal is that each of us should first identify the domain of phenomena to which he is attending and then describe its structure, that is, identify the major parameters which can be used to rank the domain elements. If we will both do this, my hunch is that we will find some degree of similarity among parameters even when we are concerned with different domains. Hopefully, these common parameters will help us bridge the communication-research gap which now stands in the way of our constructing a broad science of man. It would also be fortunate if the normal outputs of our research methodologies corresponded to the structural characteristics of our theo-

retical models. That is, a regression methodology is appropriate for a theory which emphasizes the cumulative effects of variables upon a criterion, while factorial methods are appropriate for theories which emphasize the distance relationships among variables. Fiske (1963, p. 643) has recently used this point in his plea for closer collaboration between theorists and methodologists, and Brodbeck (1963, pp. 88-91) has asked researchers to pay more attention to isomorphic relationships between model building and measurement.

The present paper deals with that broad class of personal-social events which are of concern to members of the education profession. These events have been analyzed in a number of ways, including the schemes of Bloom (1954), Krathwohl, Bloom, and Masia (1964), and Gage (1963, p. vi). My preference has been to assign such behavior to one of three domains: Personal, Interpersonal, and Attitudinal.

In the following sections an illustration is given of how descriptions of phenomena are assigned to one or another of the three domains, and each domain is conceptualized in terms of its hypothesized dimensional structure. Following Humphreys' (1962) pro-

¹ Now at San Francisco State College.

posals, two of the domains are explicated by means of facet-analysis procedures to demonstrate a desirable rational link between theory building and test construction, and factor analyses of data from 350 counselors are employed to test the empirical relationships of domain variables. Finally, interdomain correlations are used to construct a heuristic typology of counselor syndromes.

PHENOMENAL DOMAINS

The process of perceiving and describing social phenomena requires a perceiver, a stimulus object, and a verbal statement presumed to represent the cognitive outcomes of observations and introspections. The stimulus object in social perceiving may be another person, one's "self," or an interaction between persons; and the verbal statement may refer to an observable behavior, an internal behavior of the self which is not observable by others, or an inferred state or attribute. Although the last type of statement is included somewhat reluctantly because it seems to violate the phenomenal rubric, the author's personal experience has been that such descriptions are usually based on at least minimal behavioral cues, and that people can be trained to give direct attention to those cues.

Thus, a random set of social observations might include such statements as the following: (a) I see your clenched hands; (b) I feel irritable; (c) You look sad; (d) You are giving orders to him; (e) Teachers are too submissive.

In the present scheme, statements are classified in terms of the distal stimulus to which they refer. Thus, statements which refer to such non-interactive, sensory-motoric behaviors as postures, movements, and visceral states are said to belong to the Personal

Domain. In the sample of statements given above, the first is descriptive of extraorganismic behavior, the second, of intraorganismic behavior, and the third of an implied behavioral state in the other person.

Statements which refer to behavioral interactions or relationships between persons such as giving, taking, protecting, and criticizing are said to belong to the Interpersonal Domain. The fourth statement above is of this type.

Repeated encounters between a person and his environment tend to lead to a set of rather stable expectations and evaluations of that environment. These cognitive-evaluative dispositions of the person function to stabilize his perceptions of the world and others' perceptions of him and may be referred to as his expectations, sets, values, roles, and the like. Statements, like the last mentioned above, which refer to such generalizations regarding social relationships are said to belong to the Attitude Domain.

The Interpersonal Domain

The problem of the structure of the interpersonal relationship has been approached by research workers in many different fields. In spite of relative independence of purposes, traditions, and methodologies, the results of the major studies have been strikingly similar, particularly those which have used factor-analytic methods of analysis.

In a review of the investigations by Carter (1954), Leary (1957), Borgatta, Cottrell, and Mann (1958), and Schaefer (1959), it was concluded by Foa (1961) that the common findings in all studies were two substantive dimensions of interpersonal behavior which could be identified as the Dominance-Submission dimension and the Hostility-Affection dimension. Other factors appeared to be related to the

specific methodology used, to situational influences, or to sequences of actions. More recently, Adams (1964) has claimed that all interpersonal behavior, both adaptive and maladaptive, can be meaningfully categorized within these two dimensions, and he has cited additional support in the factor-analytic studies of the MMPI performed by Welsh (1956) and Jackson and Mesnick (1961).

It was also concluded by Foa (1961) that the arrangement of variables in all studies tended toward that patterned relationship known as the *circumplex* (Guttman, 1954a). That is, when the factor loadings of the variables are plotted against the two major axes, a circular order tends to appear. This circularity was most apparent among the items which comprised the Interpersonal Check List (ICL), the instrument devised by LaForge and Suczek (1955) to give operational meaning to Leary's (1957) interpersonal system of personality. Leary's system is based on the two dimensions of behavior described above, and the circularity of the ICL variables appears to be due to their having been selected rationally on this basis.

Interpersonal Space. In Figure 1 an attempt has been made to visualize the two major dimensions of interpersonal behavior in terms of a two-dimensional spatial model which corresponds in appearance to the two-factor solution of a factor-analytic procedure. Figure 1 has been marked out in sectors for heuristic purposes, that is, to lead the imagination of the reader easily to the circumplex model which will be presented in Figures 4 and 5. The dimension of emotion has been characterized by *affectionate* behaviors in one sector of the model and *critical* behaviors in the opposite sector. Interpersonal power is described as

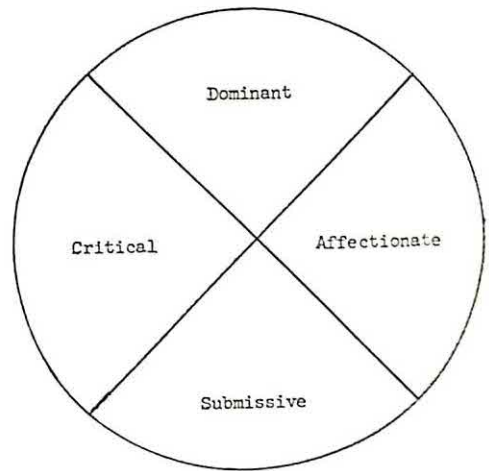


FIG. 1. Interpersonal behaviors: emotion and power.

varying from *dominant* behaviors to *submissive* behaviors.

The Attitude Domain

An attitude is identified by a consistency of response to a class of objects (English & English, 1958, p. 50). Attitudes in the field of education might have as their objects abstract social values, broad educational goals, differential curriculum purposes, course outcomes, instructional procedures, or various kinds of person-to-person interactions such as teacher-pupil relationships. One way of categorizing these objects is in terms of those having to do with goals or outcomes and those having to do with procedures or processes. It has been suggested that a general predisposition exists such that some persons tend to perceive social situations in goal-oriented terms and others in process-oriented terms, and persons with different dispositions may have special communication difficulties (Rinn, 1961a). Attitudes toward interpersonal relationships may be expected to have the same dimensionality as the actual descriptions of those relationships. For the sake of keeping a

clear distinction between the two domains, however, the attitude dimensions will be referred to by the terms *emotionality* and *control*.

Although emotionality in its pure form consists of one person liking or disliking another, emotion in an educational setting is usually expressed as an aspect of the teacher's reaction to task accomplishment or rule adherence. Thus the emotional component of a teacher's attitude toward the work of students is likely to be referred to as a tendency to give *praise* or *blame*, while the affective reactions to classroom behavior are likely to be described as the teacher's tendencies to *reward* or *punish* or to give *approval* or *disapproval*. Experimental manipulation of this variable has been described by Kounin, Gump, and Ryan (1961) in a comparison of *threatening* versus *supportive* desist techniques. Emotionality would also seem to have been a major criteria for distinguishing between *punitive* and *nonpunitive* teachers in a study by Kounin and Gump (1961).

The control dimension is the major component of such conceptual dichotomies as *traditional* versus *progressive*, *autocratic* versus *democratic* (Lippitt & White, 1958), *dominative* versus *integrative* (Anderson & Brewer, 1946), and *conservative* versus *permissive* (Lazarsfeld & Thielens, 1958). Counselors went through a long period of dispute concerning the issue of *directive* versus *nondirective* counseling (e.g., Rogers & Skinner, 1956).

Differentiation of the two ends of this continuum is not a simple process because the definition may focus on different levels and objects. At the broadest value level, there exist the contrasting doctrines of individual rights and the rights of society, social liberalism and social conservatism, the inherent wisdom of the child and

the experiential wisdom of the adult, the goodness of man and the badness of man. At the level of learning theory, belief in highly structured learning programs contrasts with belief in opportunistic development of immediate interests. At the behavioral or leadership level, teachers and counselors can give many directions or few directions, make many decisions or few decisions, and adherence to suggestions may be made a matter of rigidity or flexibility. Said another way, students may be expected to follow orders or to take initiative.

Attitude Space. As with the Interpersonal Domain, the two major dimensions of the Attitude Domain may be given spatial representation as in Figure 2. On the emotionality dimension the term *sociable* implies both warmth and closeness between persons while the term *detached* is intended to include both cool detachment and active antagonism. For the control dimension the term *permissive* suggests that the locus of responsibility of decision making has been delegated to the student and implies that the elder has faith or trust in the student's ability to benefit from taking such responsibility. The term *directive*, on the other hand, suggests that control remains in the hands of the one to whom legal responsibility has been assigned and implies that the elder is trusted to exercise this control effectively and benevolently.

The Personal Domain

What is the structure of personal behavior? A person thinks, feels, moves. The authors of the handbooks in the *Taxonomy of Educational Objectives* series (Bloom, 1954; Krathwohl et al., 1964) found that most educational objectives could be placed rather easily into one of three major domains: cognitive, affective, and psy-

chomoter. There were few references found in the literature to the psychomotor domain, and the question was raised as to whether a human being ever does acting without feeling or thinking.

Schachtel (1959) has contributed an analysis of man's sensory experiences which calls for application to the present gap between theories of perceptual, psychoanalytic, and educational development. Schachtel identifies two basic perceptual modes which he calls the autocentric and the allocentric modes of perception. The autocentric mode is characterized by the sensory experiences of pleasure or displeasure, comfort or discomfort. The allocentric mode is characterized by objectification, the phenomena of one's encounter with more or less definite environmental objects. This encounter may vary from an actively initiated exploration of the environment to a passive reception of environmental stimulation. Ginzberg (1956) has proposed a basic personality predisposition towards activity or passivity which appears to correspond to the allocentric dimension of perception.

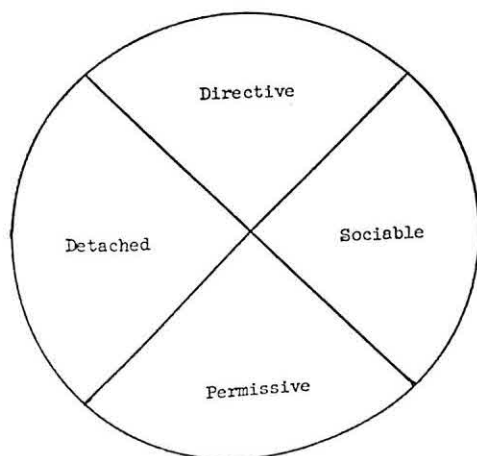


FIG. 2. Cognitive attitudes: emotionality and control.

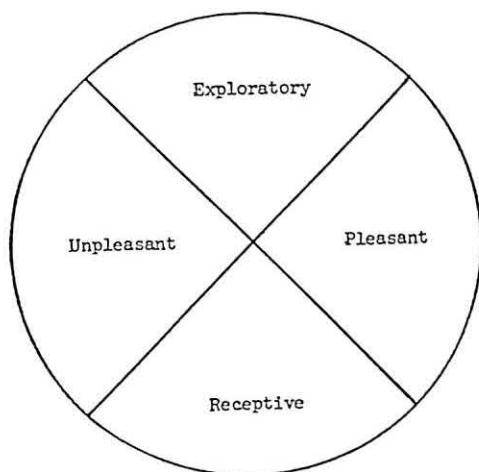


FIG. 3. Personal modes: feeling and activity.

Although I have no empirical evidence which bears directly upon the structure of the Personal Domain, I find that I am attracted to the notion that similar structures may obtain among all three domains. Therefore, I propose on a purely rational basis that the Personal Domain will be found to be structured in terms of two major dimensions, *feeling* and *activity*, which correspond to the dimensions described in the Interpersonal and Attitude domains. The feeling axis will vary from a *comfortable* state of the person on the one hand to an *uncomfortable* state on the other. The activity axis will vary from active *exploratory* behaviors to passive *receptive* behaviors, and these behaviors may be either motoric or cognitive in nature.

Since a person responds as a whole organism whenever he does respond, any action may be accompanied by varying degrees of feeling tone, from the most positive to the most negative. It seems reasonable, then, to represent the two axes as in Figure 3, where a given behavior is given spatial representation with reference to both the feeling and activity dimensions.

FACET ANALYSIS

The intellectual process of analyzing complex concepts into simple sets of elements is a logical procedure which is basic to all scientific endeavor. The term facet theory, however, has come to stand for a set of notions put forth by Guttman (1958a) regarding the structural design of social and psychological theories which utilize composite (complex) concepts. Guttman has defined composite concepts in terms of Cartesian products of the elements of simpler concepts called *facets*. Guttman (1958b) regards Guilford's (1956) work on the structure of intellect as essentially a search for more refined facets for mental tests or for the expression of gross facets of intellect as Cartesian products. Examples of facets in the area of mental ability are kind of content and level of complexity.

An important postulate of facet theory is the *contiguity principle* which states that the correlation between two variables increases with the similarity of their facet structures. Foa (1958) has applied this principle to the prediction of correlations among dyadic relationships and has demonstrated the feasibility of performing facet analyses of social relationships. Facet elements which Foa (1963) has recently identified include distinctions between observers, actors, content of behavior, object of behavior, mode of behavior, alias (point of view), and level (actual-ideal).

Facets of the Interpersonal Domain

In his review of studies of the structure of interpersonal behavior Foa (1961) proposed that the two axes of Dominance-Submission and Love-Hostility were sufficient for describing the empirical findings, but that they were not sufficient for explaining the

underlying structure which permitted a circumplex pattern to appear. The desired explanation of circularity was found in the theory of principal components of scalable attitudes developed by Guttman (1954b) based on an integration of mathematical and psychological concepts. In terms of this theory, the circumplex pattern is primarily a function of the second and fourth of the first four principal components. Foa therefore developed a fourfold facet structure of the interpersonal act based on the *content* of the action (acceptance or rejection), the *intensity* of the action (high or low), the *object* of the action (self or other), and the *mode* of the action (social or emotional). In combining these elements into profiles, Foa ignored the intensity component and redefined love and hostility as acceptance or rejection of affect (emotional mode), while dominance and submission were defined as acceptance or rejection of status (social mode). His addition of the object element (self or other) made eight facet profiles possible, and these facet combinations are listed below. It should be noted that their order is different from that proposed by Foa (1961, p. 348), and I shall say more about this later.

- A. Accept the self's power.
- B. Accept the self's emotion.
- C. Accept the other's emotion.
- D. Accept the other's power.
- E. Reject the self's power.
- F. Reject the self's emotion.
- G. Reject the other's emotion.
- H. Reject the other's power.

A unit of interpersonal behavior may now be defined in terms of a set of values for these eight facet profiles. The unit is a complex concept since it includes the relative positions of both self and other with regard to power and emotion. Thus, in a single act,

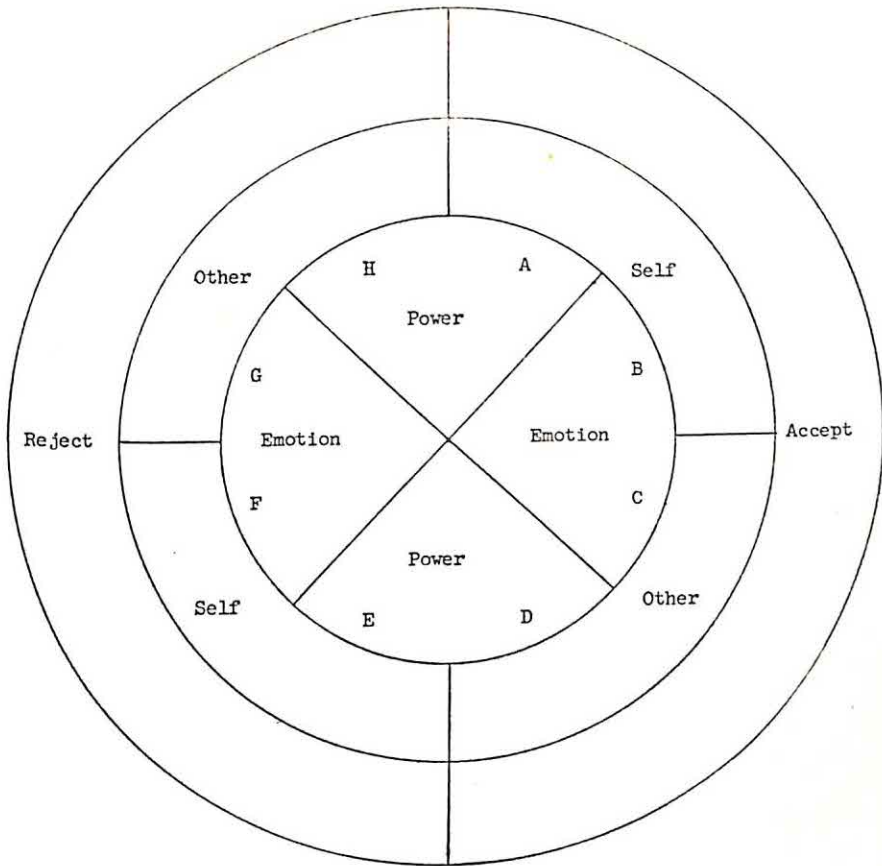


FIG. 4. Circumplex of interpersonal facets.

a person usually asserts the relative status positions of himself and another person and evokes a particular emotional relationship between himself and another.

The circular order which appears among measurement variables is assumed to depend on circularity of facets. The resolution of the order is accomplished by means of theoretical assumptions regarding facet contiguity. Visualization of the assumed facet relationships has been attempted in Figure 4 where concentric circles stand for the various facets, and sectors of the various circles represent facet elements. The letters A through H represent the eight facet profiles listed above and are related diagrammatically

to facet elements falling between the circle center and various points on the largest circumference. The relative position of these letters gives visual representation to the circularity assumed to hold among the facet profiles.

Given a circular set of facets and a set of variables purporting to represent the facet products, certain conditions are necessary before the appearance of a circumplex pattern among the variables can be predicted. One condition is that it be possible to arrange the variables and the facets (or facet profiles) into rows and columns of a matrix such that parallel diagonal lines will divide the matrix into sectors where all the facet elements have similar values.

TABLE 1
FACET COMPOSITION OF LEARY'S TYPES

Leary's types	Facet profiles							
	A	B	C	D	E	F	G	H
1. Managerial	+	+	0	0	0	0	+	+
2. Competitive	+	0	0	0	0	+	+	+
3. Aggressive	0	0	0	0	+	+	+	+
4. Skeptical	0	0	0	+	+	+	+	0
5. Modest	0	0	+	+	+	+	0	0
6. Docile	0	+	+	+	+	0	0	0
7. Cooperative	+	+	+	+	0	0	0	0
8. Responsible	+	+	+	0	0	0	0	+

This condition can be illustrated using Leary's (1957) eight interpersonal types as variables and the above A through H statements as facet profiles. In Table 1 only two values are permitted for each profile for the sake of simplicity. A plus sign is used to indicate presence of the profile, and a zero indicates absence. It can be seen that the first condition for circumplex structure is fulfilled.

Table 1 is interpreted as indicating that Leary's Managerial behavior is a composite of acceptance of self and rejection of other with regard to both power and emotion. Competitive behavior calls for accepting power in oneself and rejecting it in the other while rejecting emotion in both self and other. Aggressive behavior calls for rejection of self and other with regard to both power and emotion. Skeptical behavior involves rejection of the self's power and emotion while accepting the other's power and rejecting his emotion.

The facet ordering illustrated above differs from that proposed by Foa (1961) in that Foa, apparently for theoretical reasons, reversed the order of the *mode* components. This rever-

sal does not affect the interpretation of the Managerial, Aggressive, Modest, or Cooperative behaviors, but it gives a peculiar slant to the other four types. Thus, Foa's interpretation of Competitive behavior would be that it involved accepting emotion and rejecting power in oneself, rather than accepting power and rejecting emotion as proposed above. Again, Foa's scheme says that Responsible behavior includes rejection of the other's emotion and acceptance of his power which does not appear to be as reasonable an explanation as the other way around.

What is claimed here is that the ordering proposed in Table 1 results in behavioral units that have more face validity than those resulting from Foa's alternative ordering. Fortunately an empirical procedure can be suggested for testing how closely a rational facet structure fits a set of variables or for comparing alternative structures such as Foa's and the present one. A description of the results of a small study using this procedure is presented below primarily for the purpose of illustrating how facet theory can be applied to these questions.

Empirical Facet Analysis of ICL Items. Items from the ICL were taken as examples of variables which represented the eight interpersonal types listed in Table 1. These items were used to test the facet structure proposed here and to compare it with that suggested by Foa. To do this, eight phrases were derived from the eight facet profiles listed above and were matched by nine practicing school psychologists against each of the first 32 items on the ICL. Each item appeared on a page above the eight phrases which were in the form: "accepts self exerting social pressure toward others," and "accepts self expressing affection toward others." The other phrases either substituted rejects for accepts or reversed the positions of the words self and others. The instructions stated that phrases should be checked only if they were implied by a subject's positive response to the item, and no restriction was placed on the number of phrases to be checked.

The number of checks for each phrase on each item was tallied, and the items were grouped in terms of their application to Leary's eight types. Each cell in Table 2, then, represents the total number of times a facet was checked by nine persons on one of the four items which applied to a particular type. Maximum possible agreement among the nine raters was thus 36, and the largest actual frequency was 32. The interpretation of this particular cell frequency, to illustrate, is that there is high rater agreement on the correspondence of meaning between Coopera-

tive inventory items and the facet phrase "accepts others expressing affection toward self."

An empirical estimate of the facet profiles underlying each of Leary's types is given by the rows in Table 2. These may be compared with the rational profiles which are given by the corresponding rows in Table 1. The goodness of the fit is obscured for this set of data by the fact that the whole set of items was biased (more precisely, the interaction between raters and items was biased) in the direction of "acceptance" or "rejection" and probably also in the direction of the facet-element interactions "accepts affection" over "accepts social pressure" and in the direction of "rejects social pressure" over "rejects affection." These conclusions come from a comparison of the column totals in Table 2 and are most easily explained by the fact that the first 32 items on the ICL were selected to represent the highest level of socially desirable phrasing.

Because of these biases, the best test of goodness of fit is to ask whether the cells which correspond to the + cells (from Table 1) have higher values than the 0 cells in the same column. When this is done, it is found that the + cells are higher than all 0 cells in the same column in 21 out of 32 cases, and that the means of the four + cells are higher than those of the four 0 cells in all eight columns.

The data presented in Table 2 also make it possible to compare alternative orderings of facets. To illustrate, Foa's (1961) pro-

TABLE 2
EMPIRICAL FACET ANALYSIS OF 32 ICL ITEMS

Leary's types	Facet profiles							
	A	B	C	D	E	F	G	H
1. Managerial	24	25	22	19	10	7	9	18
2. Competitive	24	22	22	14	11	10	11	19
3. Aggressive	29	17	18	17	5	10	11	15
4. Skeptical	20	16	14	17	10	17	20	12
5. Modest	12	18	20	28	20	11	9	4
6. Docile	14	26	30	25	17	5	5	4
7. Cooperative	17	27	32	27	9	4	0	3
8. Responsible	23	29	29	25	7	3	4	6
Totals	163	180	187	172	89	67	69	81

posal is compared with the one presented here. Since the points of difference have to do only with the profiles of Leary's even-numbered types, the comparison has been restricted to the disputed cells. The comparison is made by counting the number of hits in the disputed cells of Table 2 for each proposal. Thus Foa had proposed that the set of facet profiles for Leary's Competitive type should consist of Profiles B, E, G, and H, while the present proposal is for Profiles A, F, G, and H. If the common cells (G and H) are eliminated and the other values summed, the result is a standoff: 33 to 34. In the other three cases, however, (types 4, 6, and 8), the comparison favors the present proposal over Foa's: 37 to 26, 43 to 19, and 35 to 29 for the Skeptical, Docile, and Responsible types, respectively. Thus, although the sample of raters was small, the facet structure proposed here is supported, and an important rational link between theory building and test construction was illustrated.

The converse of this process—after determining on theoretical grounds the facet design which underlies a particular domain of investigation—would be to write items which meet the specifications of each combination of facet elements. The intercorrelations of variables thus defined could then be predicted in terms of their facet elements.

Facets of the Attitude Domain

Several considerations went into the development of a facet structure for the Attitude Domain. The first had to do with identifying the facets which might underlie educational attitudes. This was approached by asking if the facets which had been used to analyze statements about interpersonal interactions could be applied to statements regarding educational attitudes. A positive answer was found to be in harmony with the present position that (a) the educational setting is one in which certain kinds of interpersonal interactions occur, (b) teachers and counselors take value positions with regard to the kind of interactions which they believe ought to occur, (c) educators are concerned with the issue of the relative degree of status, control, or respect which ought to be held by

teachers, students, and others, and (d) educators are concerned with the kind and degree of emotional relationships which are likely to contribute to the achievement of various educational goals. The school-counselor population in particular can be expected to be sensitive to the interpersonal elements of educational practices.

On the basis of the above assumptions, then, it was held that the facets which underlie educational attitudes should be similar to those which were used to explicate the varieties of interpersonal behavior. It will be recalled that in the case of interpersonal behavior, the first of these facets had to do with the degree of acceptance or rejection which the subject gave to some aspect of the social situation (Foa's *content* facet). The social object was assumed to be a person, either oneself or another (the *object* facet), and the significant interpersonal characteristics of persons were their power relationships and their emotional relationships (the *mode* facet).

The first facet of educational attitudes is provided by the degree of *favorability* which the subject feels towards various educational practices, the extent to which he accepts or rejects a particular policy. The concept of *favorability* is used by Guttman (1954b) in his analysis of the principal components of attitudes, and it corresponds to Foa's *content* facet in the Interpersonal Domain.

The second facet in the Interpersonal Domain has been assumed to be the object of the action (self or other), but in the process of identifying the object of an attitude a question arose. Is it more appropriate to consider that the favorability of an educational attitude (acceptance or rejection) is directed toward a person (self or other) or toward a mode of behavior (social or emotional)? My assumption is that

educators tend to be sensitive primarily to the overall interpersonal climate of the school and that they are secondarily concerned with the particular persons who may have contributed to the climate. Therefore, the second attitude facet was identified as the mode (emotionality or control) of the attitude, retaining the title used in the Interpersonal Domain.

Having assigned the *mode* component as the second attitude facet, the third facet was identified as the *actor* (self or other). This new term was used as a reminder that it was now the behavioral *mode* which was seen as the object of attention and not a person. The distinction might be characterized by saying that in the case of inter-

personal action, the subject directs his evaluation toward some person with a certain characteristic; in the case of educational practice, the subject directs his evaluation toward a social interaction to which some person contributes.

Having identified the facets purported to underlie educational attitudes, the eight facet products or profiles can be listed as follows:

- A. Acceptance of control in self.
- D. Acceptance of control in other.
- C. Acceptance of emotion in other.
- B. Acceptance of emotion in self.
- E. Rejection of control in self.
- H. Rejection of control in other.
- G. Rejection of emotion in other.
- F. Rejection of emotion in self.

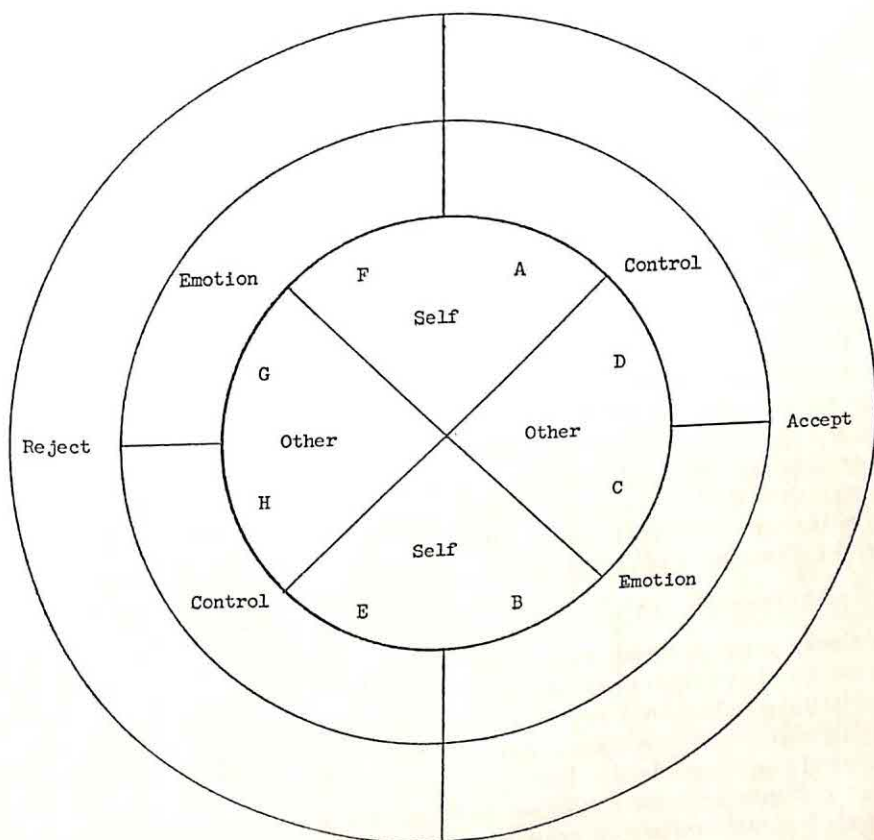


FIG. 5. Circumplex of attitude facets.

It should be noted that the identifying letter was taken from that profile in the interpersonal-behavior domain to which this product most closely corresponds. Thus, "acceptance of control in self" appeared to relate most closely to "accepts the self's power," and "acceptance of control in other" to "accepts the other's power." Although these products are not claimed to be exact equivalents, they form the basis upon which we can predict similar responses on instruments used to measure the two domains.

An attitude unit may now be defined as the product of two or more of the eight facet profiles listed above. A single attitude usually contains the subject's relative preferences for control and emotionality in both himself and others. Since the facets are assumed to have a circular relationship, they can be represented visually as in Figure 5 in a manner similar to that of the Interpersonal Domain.

A comparison of the facet structures of the Interpersonal and Attitude domains can now be made. In the Interpersonal Domain, the eight profiles are defined as the product of three facets: the *content* of the action (accept or reject), the *object* of the action (self or other), and the *mode* of the action (emotion or power). In the Attitude Domain, the three facets are the *favorability* of the attitude (acceptance or rejection), the *mode* of the attitude (emotionality or control), and the *actor* in the situation (self or other).

FACTOR ANALYSIS

The three most common objectives of factor-analytic methodology are parsimony, orthogonality, and psychological meaningfulness. The most controversial of these issues is the last. In a recent critique of issues relating to the psychological interpretation of factors, Coan (1964) has described the

various positions which have been taken with regard to the relationship between factors and reality. At the minimum, a factor is a statistical function of a pattern of correlations among psychological variables, and it may therefore be used as a convenient principle of classification. Depending on the variables used, it may also be accorded the status of an underlying trait lying outside the range of observation. Traits may be spoken of as theoretical constructs without reference to social or biological determination, or they may be assumed to represent fundamental functions of the organism, perhaps at the neurological level. At a high level of generality the term *type* may be used to indicate a syndrome of attributes or trait indicators.

In the present study factor analysis provides a methodology for testing the structure of a set of variables against that predicted on the basis of a facet analysis of the relevant phenomenal domain. In the following sections two standard psychological inventories are analyzed to determine how closely their empirical factor structures correspond to the dimensionality of their respective theoretical models.

Interpersonal Attributes of Counselors

As a basic instrument for measuring the domain of interpersonal phenomena in a population of teachers or counselors, LaForge and Suczek's (1955) Interpersonal Check List (ICL) has several useful characteristics. As has been discussed above, the instrument comes closer than most to meeting the criteria of exhaustiveness in the Interpersonal Domain on both theoretical and empirical grounds. Although Bales (1951) has also set up his category system to have general applicability for face-to-face interactions, his system is usually adapted for use as a sys-

tematic observation technique and does not lend itself well to self-description.

Another important characteristic of the ICL is its objectivity. That is, it is composed of items which tend to describe the overt or public manifestations of personality rather than the covert or private level (Leary, 1957, pp. 132-136). It therefore lends itself to use as a research tool in situations where it is appropriate or strategic to collect information at the public level of behavior, and in a theoretical context which is built upon the perceivable aspects of personal relationships rather than on constructs presumed to operate at the unconscious level of personality. The items also tend to be descriptive of the normal range of behaviors rather than the abnormal, as in the MMPI and other clinically based instruments, so that the inventory can be used with self-descriptive instructions with minimal arousal of feelings of threat. These characteristics combine to make the ICL a research instrument which is appropriate for use in educational settings and is compatible with a theoretical framework of personality that emphasizes the role of cognitive awareness of self in the determination of human behavior.

Collection of Data. A list of all fully certified Ohio secondary-school guidance counselors (approximately 500) was secured from the Ohio State Department of Guid-

ance.² Each person on the list was sent a packet of materials consisting of a covering letter, a copy of the ICL marked "Self," a copy of the ICL marked "Ideal Counselor," a copy of the Teacher Preference Schedule (TPS)—to be described later—, a TPS answer sheet, a sheet of instructions for completing the three inventories, and a stamped, addressed return envelope. Six weeks later a follow-up letter was sent to each person who had not yet returned the inventories.

At the end of another month, over 400 sets of materials had been returned, and of these, 350 were complete and usable for all three inventories. The responses of these 350 counselors comprised the basic data for this study (Rinn, 1961b). Although detailed analysis of these data will be reported elsewhere, some of the findings are relevant to the present discussion of the structure of the Interpersonal Domain and are presented in the following sections.

Factor Analysis of ICL Conventional Scores. When conventional scoring procedures are used on the ICL, eight scores are derived by summing up responses in each area of psychological content (the interpersonal traits) across four "levels of favorability" (Leary's "intensity levels"), and the items in each content area are weighted equally regardless of favorability level. The score intercorrelations for the present data are presented in Table 3. A principal-components factor solution of this matrix produced the circular arrangement of scores shown in Figure 6, which can be seen to deviate only slightly from that predicted. That is, on the basis of the dimensional

² This part of the study was supported by a National Defense Education Act grant to the State of Ohio Department of Education while the author was an assistant professor of education at the Ohio State University.

TABLE 3
INTERCORRELATIONS OF ICL CONVENTIONAL SCORES

	1	2	3	4	5	6	7	8
1								
2	61							
3	55	67						
4	38	51	64					
5	07	09	19	52				
6	13	11	16	32	68			
7	31	03	00	01	35	52		
8	29	-02	00	03	37	46	72	

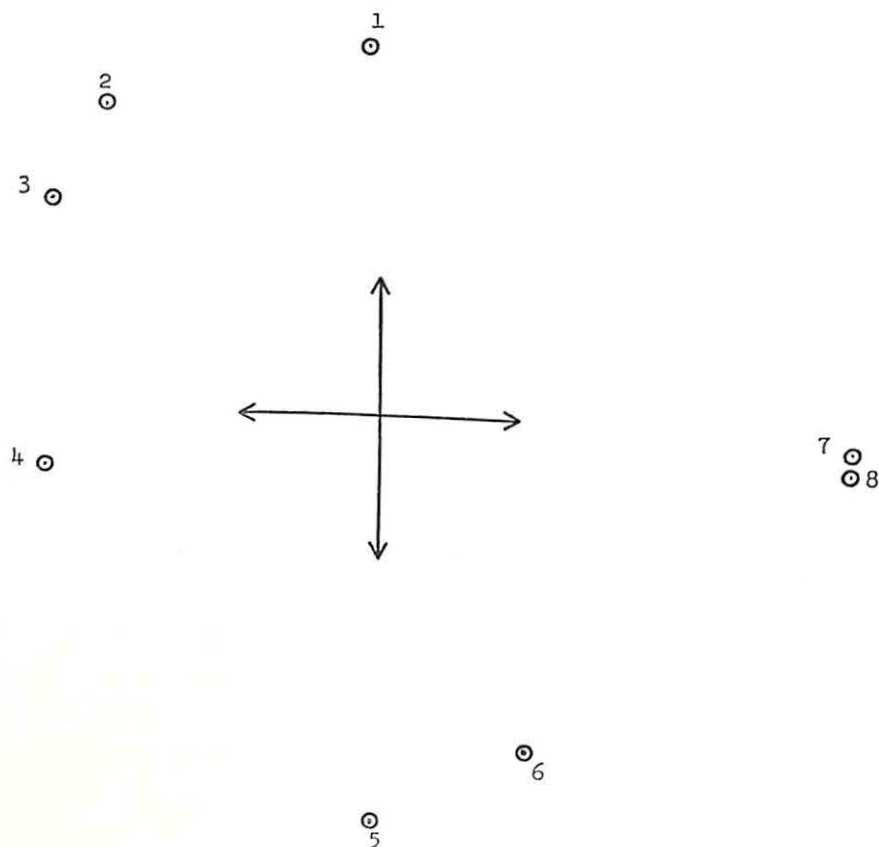


FIG. 6. Factor plot of ICL conventional scores.

analysis of the Interpersonal Domain shown in Figure 1 and the facet analysis of Leary's types shown in Table 1, it was predicted that a circumplex structure would obtain among variables which sampled the Interpersonal Domain. This prediction is clearly supported by the plot of ICL scores shown in Figure 6.

The analysis also produced a general factor upon which all eight scales had approximately equal loadings. It was interpreted as an "elevation" factor which was due to the tendency of persons to spread their responses over different portions of the ICL responses scale. When the variance due to this elevation response set was removed by equalizing the mean item elevation for all persons, a new factor analysis produced only two factors, and the pattern of scores did not change from that shown in Figure 6.

New ICL Cluster Scores. When the correlations among individual ICL items were examined, it appeared that the conventional scoring procedure had not grouped the items as efficiently as possible. Therefore, it was

decided to develop a new scoring procedure which would maximize the distance between item clusters and minimize the overlapping of items in factor space. A new cluster analysis of the individual items was performed with the aim of selecting the smallest number of independent item clusters with the highest internal reliability. This analysis was greatly facilitated by the cluster-analysis computer programs recently developed on the Berkeley campus of the University of California under the direction of Robert Tryon.

The outcome of this procedure was a set of 4 scores derived from 4 clusters of 10 items each which were identified by the names Dom, Hos, Sub, and Lov.³ These 4 clusters of items are encircled on the factor plot shown in Figure 7. This plot shows the relative positions of the first 64

³ ICL item numbers are Dom: 3, 4, 6, 35, 36, 37, 38, 39, 42, 44; Hos: 8, 12, 13, 14, 15, 41, 43, 45, 46, 47; Sub: 18, 20, 22, 24, 26, 50, 51, 55, 56, 58; Lov: 27, 28, 29, 31, 32, 57, 59, 60, 61, 64.

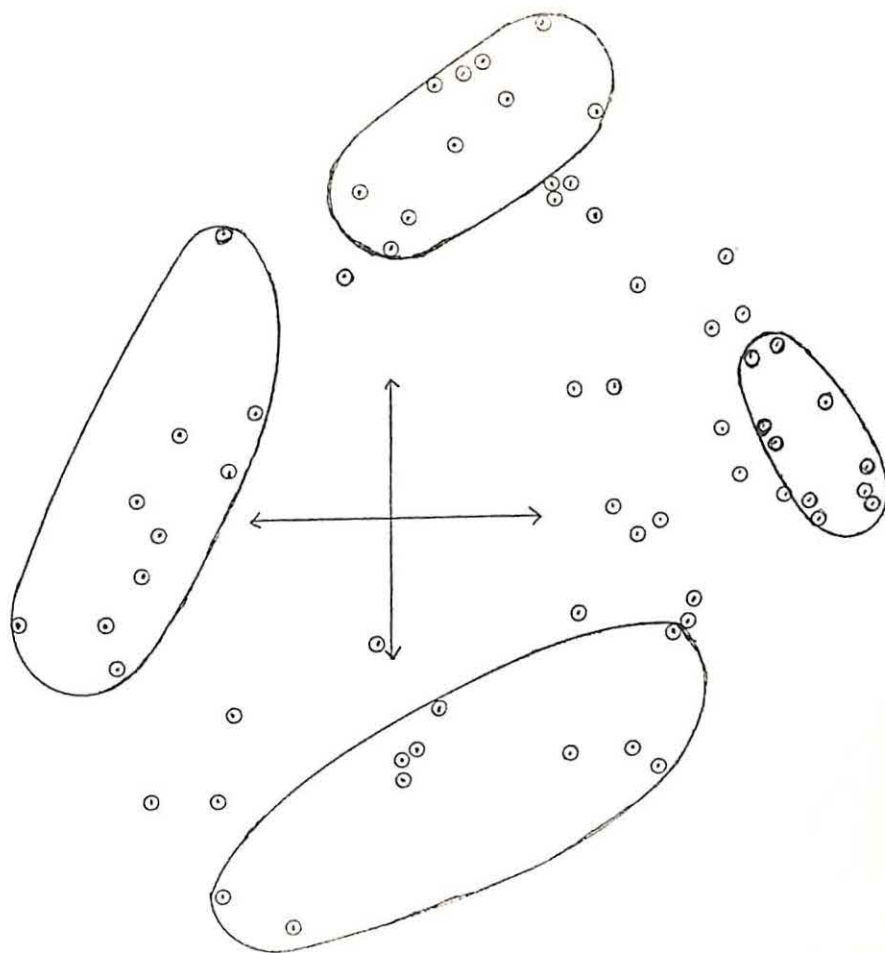


FIG. 7. Factor plot of ICL items 1-64.

items on the ICL from which all 40 items were selected. The reliability coefficients of the clusters are satisfactorily high (.87, .77, .81, .89, respectively) even though each cluster contains only 10 items. The cluster intercorrelations show a high degree of orthogonality and are reported in Table 4. The four item clusters are empirical representations of the four sectors of the inter-person circumplex diagrammed in Figure 1.

Educational Attitudes of Counselors

The Teacher Preference Schedule (TPS) is the product of an investigation by Stern, Masling, Denton, Henderson, and Levin (1960) into the nature of educational attitudes and gratifications. It has its theoretical base in the contention of Stern, Stein,

and Bloom (1956) that unconscious personal motives contribute to teacher effectiveness and satisfaction. Thus, the study of motives for teaching may be seen as part of the broader investigation into the nature of motives which

TABLE 4
INTERCORRELATIONS OF NEW ICL
CLUSTER SCORES

	Dom	Hos	Sub	Lov
Dom				
Hos	.07			
Sub	-.09	.15		
Lov	.37	-.27	.34	

underlie career choice and development.

The use of the TPS is appropriate for the present study not only because secondary-school counselors come from the teaching ranks and can be presumed to share at least some of the common motives of teachers, but also because the guidance function is inextricably linked to the general objectives of education. Thus, although the motives of guidance counselors might be expected to differ in some ways from those of teachers (indeed, the nature of these differences needs to be investigated), it is quite appropriate and desirable to seek the expression of these motives within the general framework of educational goals and practices.

Two instruments were developed by Stern et al. (1960) for the separate measurement of teaching gratifications (Form G) and teaching attitudes (Form A). The latter (Form A958) was selected for the present study as being an appropriate measure of the Attitude Domain to use with school counselors. In the process of instrument development, 10 different patterns or "roles" were identified and are described in terms of their gratification and attitude components in Table 5. Although Stern et al. did not specify

any ordered relationship among the 10 scales, the roles seemed to fit quite nicely into the dimensional structure hypothesized for the Attitude Domain, especially the pair labeled Dominant and Dependent and the ones called Critical and Nurturant.

Factor Analysis of the TPS Scales. The items in the TPS scales were originally derived from statements made by actual teachers who had been identified as prototypes of the teaching roles. Significant item discriminations for all scales were reported (Stern et al., 1960) based on three samples of teachers, practice teachers, and teacher trainees.

Stern et al. do not appear to have investigated the interscale relations in a systematic way, but they did report that simple inspection of tetrachoric correlation matrices revealed two clusters of scales. Cluster 1 consisted of the Exhibitionistic, Nurturant, Preadult, and Nondirective scales and was labeled *child centered*. Cluster 2 consisted of the Practical, Dominant, Orderly, and Dependent scales and was labeled *teacher centered*. The other two scales (Critical, Status-striving) correlated with both clusters, and the interpretation was offered that these two attitudes might have different manifestations depending on whether

TABLE 5
ANALYTIC SUMMARY OF TPS SCALES

Role	Gratifications	Attitudes
1. Practical	Instrumental rewards	Detachment
2. Status-striving	Prestige	Professional dignity
3. Nurturant	Children's affection	Providing love
4. Nondirective	Children's autonomy	Encouraging self-actualization
5. Critical	Promoting teachers' rights	Reforming schools
6. Preadult	Vicarious participation	Identification with children
7. Orderly	Obsessive activities	Developing good pupil habits
8. Dependent	Support from superiors	Cooperation with authority
9. Exhibitionistic	Children's admiration	Showmanship
10. Dominant	Children's obedience	Maintaining discipline

TABLE 6
INTERCORRELATIONS OF TPS SCALES

	1	2	3	4	5	6	7	8	9	10
1										
2	-.09									
3	-.09	.47								
4	-.12	.33	.53							
5	.48	-.09	.06	.06						
6	.13	.28	.54	.38	.35					
7	.39	.29	.26	.03	.35	.30				
8	.15	.48	.43	.24	.00	.34	.51			
9	.24	.15	.35	.25	.39	.45	.36	.23		
10	.56	.05	-.01	-.26	.45	.21	.60	.28	.32	

they were associated with a child-centered or a teacher-centered orientation. In terms of overall mean differences between the scales, there was reported a consistent tendency for the Cluster 1 scales to receive the higher means and the Cluster 2 scales the lower.

To begin the investigation of the interscale relationships for the present sample of 350 counselors, the correlation matrix was calculated and is presented in Table 6. Inspection of this

matrix showed several obvious discrepancies from the two-cluster pattern reported by Stern et al. (1960). Since the matrix utilized the Pearson correlation coefficient and was based on a larger sample of persons than the original study, it was feasible to perform a standard factor analysis of the 10 scales.

A principal-components solution yielded only three significant factors by the criteria of abrupt slope change on the eigenvalue curve (eigenvalues of

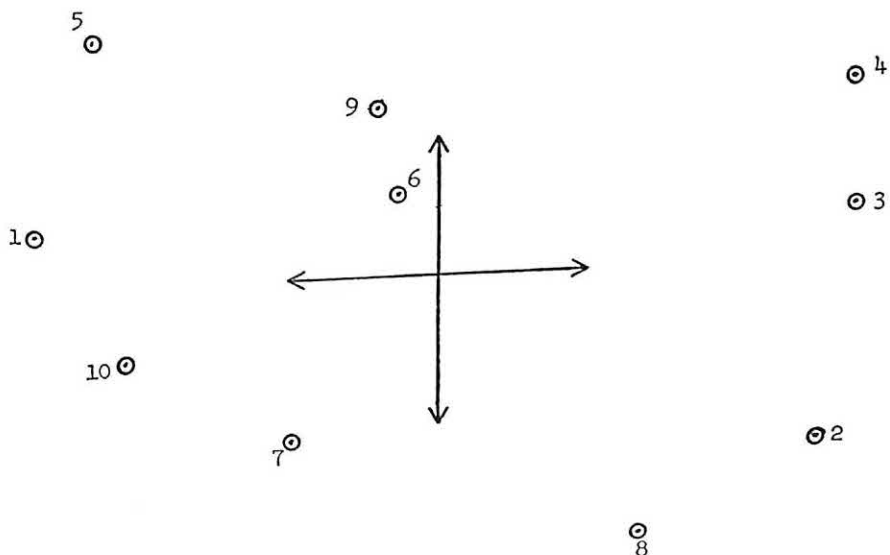


FIG. 8. Factor plot of TPS scales.

3.39, 1.73, and .53). The first factor loaded approximately equally on all 10 scales, corresponding to the general factor in the ICL analysis. It was similarly interpreted as primarily due to the elevation variance, that is, to the tendency among persons to use a particular range on the answer sheet when responding to inventory items. A plot of the loadings of the 10 scales on the second and third factors is shown in Figure 8. In addition to modifying the general picture given by Stern et al., this analysis revealed that the TPS scales were related in a quasi-circumplex pattern similar to that found to underlie the ICL.

Since the TPS factor pattern was similar to that of the attitude circumplex which had been developed on theoretical grounds, its dimensionality was examined for correspondence by means of a rational analysis of the TPS scales and items. This analysis revealed several interesting characteristics of the TPS scales which had not been reported by Stern et al. (1960).

Although most TPS items were found to refer to only two actors (an initiator and a receiver), the full set of items included three classes of persons: teachers, students, and administrators (the latter category including school boards and parent groups). This meant that the interpersonal-relations component of statements might refer to either of two dyads: teacher-student relations and teacher-administrator relations. Furthermore, the items which describe the two types of relations were not spread evenly throughout the inventory but were concentrated in various scales. Thus, Scales 3, 4, 7, 10 refer to teacher-student relations and Scales 1, 5, 2, 8, to teacher-administrator relations. Scales 6 and 9 referred mainly to nonrelational, internal states of the teacher and posed a special problem that is discussed below.

With regard to the quality of the interpersonal relationships, the horizontal axis in Figure 8 was found to correspond to the *emotionality* dimension hypothesized for the Attitude Domain. Thus, Scales 3, 4, 2, 8 reflected sociable or affectionate relations, and Scales 1, 5, 7, 10 reflected critical or detached relations.

Upon analyzing the scales along the vertical axis of Figure 8, it was noticed that at one pole Scales 3 and 4 supported the rights of students to express themselves, to make decisions and the like, and Scales 1 and 5 argued for the rights of teachers in matters of policy making, school reform, and personal independence. At the other pole, Scales 7 and 10 denied student independence and initiative, and Scales 2 and 8 supported teacher dependence upon and cooperation with the authority structure of the school. All of this suggested that the vertical axis in Figure 8 reflected the concept of *personal autonomy* wherein relative independence of action was contrasted with conformity to rules and authorities. Although the concept of *autonomy* is clearly related to the hypothetical dimension of *control*, it appears in a complex form in the TPS. That is, it appears on various scales in the form of teacher independence of administrators (Scales 1 and 5), student independence of teachers (Scales 3 and 4), freedom for the teacher to be a showman (Scale 9), and freedom for the child to be a child (Scale 6).

In terms of the above analysis, an advocate of Stern et al.'s (1960) "teacher-centered" syndrome would favor teacher autonomy and student obedience, while the "child-centered" teacher would support student autonomy but give obedience to his own superiors.

TPS Cluster Scores. Since item analyses of the 10 TPS scales indicated

TABLE 7
CORRELATIONS BETWEEN ICL CLUSTER SCORES AND TPS SCALES

ICL scores	TPS scales									
	Crt 5	Pra 1	Dom 10	Ord 7	Dep 8	Sta 2	Nur 3	Non 4	Exh 9	Pre 6
Dom	-.03	-.06	.02	.19	.15	.17	.03	-.02	-.02	-.04
Hos	.05	.08	.12	.01	-.07	-.08	-.11	-.06	.00	-.08
Sub	.07	.08	.24	.25	.19	.12	.17	-.05	.15	.16
Lov	.06	-.02	.00	.27	.25	.20	.28	.10	.14	.15

Note.— $r = .14$ is significant at the .01 level.

that internal reliabilities and relative independence of clusters were comparable to those reported in the original study, and since the rational analyses described above showed that the various scales had different facet structures with regard to the *autonomy* dimension, it was decided to retain the original scoring formulas rather than to form quadrant scores as was done with the ICL. Therefore, 10 cluster scores were calculated for each counselor and were identified by means of the original TPS numbers and names listed in Table 5.

DOMAIN RELATIONSHIPS

Three events now led the author to expect certain relationships to obtain between the measures of the Attitude and Interpersonal domains. First, the two domains were described in terms of the same set of facets, and the facet elements in each case were related in terms of a circular order. Second, the responses to the measures of each domain were such as to produce complex structures of variables which appeared to correspond to the theoretical designs of the domains. The author's own need for consonance led to the thought that a counselor who engaged in, say, "dominative" interpersonal behaviors would subscribe to a "directive" educational philosophy, and that a counselor who engaged in "affection-

ate" behaviors would prefer a "sociable" approach to teaching, and so on.

Thus, the pairs of scores from different instruments having similar structures were examined with the expectation that scores derived from similar sectors of the respective circumplexes would covary positively and that scores derived from complementary sectors would covary negatively. These expectancies were tested by correlating the four ICL cluster scores of the 350 counselors with their 10 TPS scale scores. These correlations are presented in Table 7 where the TPS scales are reordered in terms of the circular factor pattern shown in Figure 8.

Inspection of Table 7 showed that the correlations in general were not as high as might be expected on purely rational grounds, and that more of the existing correlation was related to the axes representing emotion and emotionality than to those representing interpersonal power and attitudinal autonomy. To some extent this finding may be attributed to the complex construction of the TPS scales along the latter dimension, but it seems also to suggest that counselors feel more "of one piece" with regard to their affective characteristics and relationships than they do with regard to their influence relationships. That is, it is hypothesized that counselors as a

group have more conflict and anxiety associated with the question of how much control to exert over students and colleagues than they do with the question of what is appropriate emotional expression. The author's experience with counselors in training tends to agree with this hypothesis, and it can also be pointed out that more heat has been generated over the directive-nondirective issue in counseling than any other.

Counselor Types

A final characteristic of the phenomenal model which has not yet been discussed is the construction of phenomenal syndromes, profiles, or types. Following Guilford's (1959, pp. 99-106) hierarchical conception of personality structure (cf. Coan, 1964, p. 137), a syndrome type is identified by a pattern of positive correlations among a set of distinct attributes or traits. It is suggested here that the phenomenal syndrome should contain attributes from two or more domains and that this procedure will permit the identification of congruent and incongruent personality profiles. Thus, an example of a congruent profile would be the counselor who takes a dominant interpersonal role and who justifies his behavior in terms of a directive educational philosophy or attitude. In contrast, an incongruent counselor might be one who behaves dominantly but who claims to believe in permissive educational procedures.

Because of the fact that the two instruments selected to represent the Interpersonal (ICL) and Attitude (TPS) domains proved not to have completely parallel structures, it is not possible to give empirical support at the present time for the full set of hypothetical profiles. However, the construction of empirical profiles can be illustrated by utilizing the patterns

of correlation which actually exist among the present variables. When each ICL score in Table 7 was examined in terms of its significant correlations with TPS scales, it was possible to construct four heuristic counselor syndromes. Since some attitudes (order, dependence) correlated positively with all ICL types (except the "critical" counselor), those elements were considered to be general effects—perhaps occupational characteristics—and were omitted from the syndromes. Each profile consists of a typical interpersonal-behavior pattern with its supporting attitudinal rationalizations. The actual phrases are taken or adapted from items in the ICL and TPS instruments, and since these are self-report inventories, the profiles are stated in the first person.

Managerial Counselor: I am self-confident and assertive, and I guide and direct people. They admire and defer to me because I maintain professional dignity and propriety.

Detached Counselor: I am skeptical and sometimes indifferent, and I can be critical and irritable. I get little gratification from my work.

Conforming Counselor: I admire and imitate others, and I want their advice and approval. Children cheerfully obey me because I give them loving discipline.

Sociable Counselor: I am always friendly and agreeable, and I help and encourage others. Children hold me in esteem because I give them affection without losing professional dignity.

REFERENCES

- ADAMS, H. Mental illness or interpersonal behavior. *American Psychologist*, 1964, 19, 191-197.
- ANDERSON, H. H., & BREWER, J. E. Studies of teachers' classroom personalities: II. Effect of teachers' domineering and integrative contacts on children's classroom be-

- havior. *Applied Psychological Monographs*, 1946, No. 8. P. 128.
- BALES, R. F. *Interaction process analysis*. Cambridge, Mass.: Addison-Wesley, 1951.
- BLOOM, S. (Ed.) *Taxonomy of educational objectives: Handbook I. Cognitive domain*. New York: Longmans, Green, 1954.
- BORGATTA, E. F., COTTRELL, L. S., JR., & MANN, J. M. The spectrum of individual interaction characteristics: An interdimensional analysis. *Psychological Reports*, 1958, 4, 279-319.
- BRODBECK, MAY. Logic and scientific method in research on teaching. In N. L. Gage (Ed.), *Handbook of research on teaching*. Chicago: Rand McNally, 1963. Pp. 44-93.
- CARTER, L. F. Evaluating the performance of individuals as members of small groups. *Personnel Psychology*, 1954, 7, 477-484.
- COAN, R. W. Facts, factors, and artifacts: The quest for psychological meaning. *Psychological Review*, 1964, 71, 123-140.
- ENGLISH, H., & ENGLISH, AVA. *A comprehensive dictionary of psychological and psychoanalytical terms*. New York: Longmans, Green, 1958.
- FISKE, D. Homogeneity and variation in measuring personality. *American Psychologist*, 1963, 18, 643-652.
- FOA, U. The contiguity principle in the structure of interpersonal relations. *Human Relations*, 1958, 11, 229-238.
- FOA, U. Convergences in the analysis of the structure of interpersonal behavior. *Psychological Review*, 1961, 68, 341-353.
- FOA, U. A structural theory of interpersonal behavior. Unpublished manuscript, 1963, Israel Institute of Applied Social Research.
- GAGE, N. L. *Handbook of research on teaching*. Chicago: Rand McNally, 1963.
- GINZBERG, E. *Occupational choice: An approach to a general theory*. New York: Columbia Univer. Press, 1956.
- GUILFORD, J. P. The structure of intellect. *Psychological Bulletin*, 1956, 53, 267-293.
- GUILFORD, J. P. *Personality*. New York: McGraw-Hill, 1959.
- GUTTMAN, L. A new approach to factor analysis: The radex. In P. F. Lazarsfeld (Ed.), *Mathematical thinking in the social sciences*. Glencoe, Ill.: Free Press, 1954. Pp. 258-348. (a)
- GUTTMAN, L. The principal components of scalable attitudes. In P. F. Lazarsfeld (Ed.), *Mathematical thinking in the social sciences*. Glencoe, Ill.: Free Press, 1954. Pp. 216-257. (b)
- GUTTMAN, L. Introduction to facet design and analysis. In, *Proceedings of the Fifteenth International Congress of Psychology*. Brussels, 1957. Amsterdam: North-Holland, 1958. (a)
- GUTTMAN, L. What lies ahead for factor analysis? *Educational and Psychological Measurement*, 1958, 18, 497-515. (b)
- HUMPHREYS, L. The organization of human abilities. *American Psychologist*, 1962, 17, 475-483.
- JACKSON, D. N., & MESSICK, S. Acquiescence and desirability as response determinants on the MMPI. *Educational and Psychological Measurement*, 1961, 21, 771-790.
- KOUNIN, J. S., & GUMP, P. V. The comparative influence of punitive and non-punitive teachers upon childrens' concepts of school misconduct. *Journal of Educational Psychology*, 1961, 52, 44-49.
- KOUNIN, J., GUMP, P., & RYAN, J. Explorations in classroom management. *Journal of Teacher Education*, 1961, 12, 235-246.
- KRATHWOHL, D., BLOOM, B., & MASIA, B. *Taxonomy of educational objectives: Handbook II. Affective domain*. New York: McKay, 1964.
- LAForge, R., & SUCZEK, R. The interpersonal dimension of personality: III. An interpersonal checklist. *Journal of Personality*, 1955, 24, 94-112.
- LAZARSFELD, P., & THIELENS, W. Social scientists and recent threats to academic freedom. *Social Problems*, 1958, 5, 244-266.
- LEARY, T. *Interpersonal diagnosis of personality*. New York: Ronald, 1957.
- LIPPITT, R., & WHITE, R. An experimental study of leadership and group life. In E. Maccoby, T. Newcomb, & E. Hartley (Eds.), *Readings in social psychology*. (3rd ed.) New York: Holt, Rinehart & Winston, 1958. Pp. 26-30.
- RINN, J. Group guidance: Two processes. *Personnel and Guidance Journal*, 1961, 39, 591-594. (a)
- RINN, J. Personality characteristics of 350 Ohio counselors as shown by responses on the Interpersonal Check List and the Teacher Preference Schedule. Mimeograph report, August 1961. (b)
- ROGERS, C., & SKINNER, B. Some issues concerning the control of human behavior: A symposium. *Science*, 1956, 124, 1057-1066.
- SCHACHTEL, E. *Metamorphosis: On the development of affect, perception, attention*,

Braine clearly recognizes that the ability to deal with novel sentences constitutes a major part of the child's linguistic competence. It is thus central among the phenomena that a theory of language learning must explain. Braine takes the ability to construct and comprehend novel sentences to be a special case of transfer of training based upon stimulus and response generalization. Specifically, contextual generalization is identified as a case of perceptual learning, "a process of auditory differentiation, or of becoming familiar with, the temporal positions of expressions in utterances [p. 326]."

ARGUMENTS FOR CONTEXTUAL GENERALIZATION

Braine notes the inadequacy of a theory of syntax acquisition based on associative relations between lexical items. Such a theory cannot account for the ability of speakers to recognize the grammatical structure of nonsense material. Thus, a nonsense syllable, *kivil*, is recognized as a nonce verb in:

1. People *kivil* every day.

That is, 1 is recognized as syntactically well formed despite the lack of associative connections between the words.

Such examples show that the speaker's ability to exploit syntactic relations does not depend upon forming associative bonds between lexical items. Therefore, Braine argues, the speaker's information about syntactic structure primarily concerns the grammatical properties of *locations* in sentences. Thus, the child learns such facts as: The first position in a simple English sentence is characteristically the noun position; the second position is characteristically occupied by a verb. Since the syntactic properties of a position do not, by definition, depend on the lexical item or phrase that appears in that position, novel material in a given location is perceived as having the grammatical properties characteristic of that location. Thus *kivil* is recognized as a verb in 1 because it appears in the position assigned the verbal element in such sentences as:

2. The boys eat the rabbits.
3. The boys do eat the rabbits.

For Braine, then, the syntactic properties of a segment are determined by the locations in which it occurs. A verb is defined as a word which characteristically appears in the second position in simple sentences, a noun is a word which characteristically appears in the first position in simple sentences, etc. To learn the syntactic properties of a word is primarily to learn the positions in which it can occur. In particular, given the positions a word can occupy in one sentence, we can often predict the positions which it may occupy in new sentences. From the fact that *eat* occupies the second position in 2, we can predict that it will occupy the homologous position in 4.

4. The wolves eat the rabbits.

Syntax assimilation thus consists of generalizing information about the positions in which a word is observed to occur. The correct use of a given word in a given position in new sentences is a consequence of such processes of generalization.

Braine holds that a description of word order accounts for much of the grammar of English and, consequently, that a theory which accounts for the learning of positions will have considerable explanatory power. Braine admits, however, that learning syntactic relations cannot consist solely of learning the appropriate relative positions of words and phrases. First, Braine points out, knowledge of relative positions would contribute little to the mastery of languages in which syntactic relations are expressed by inflection rather than order. Second, the notion of a position must be construed sufficiently abstractly so that a given sentential position can be occupied either by a word or by a phrase. (Notice, for example, that the "second" position in 3, i.e., the position functionally equivalent to the one occupied by "eat" in 2, is filled by the phrase "do eat.")

Some explanation is required for the fact that phrases may exhibit positional privileges analogous to those exhibited by single words. In short, an explanation is needed of how phrases can act as syntactic units.

To accommodate inflection as a syntactic device and to account for the integrity of the phrase, Braine resorts to a limited associationism. He postulates associative bonds between "closed-class" morphemes (e.g., inflections) and "open-class" morphemes such as nouns and verbs. For example, Braine would presumably hold that the phoneme "s" at the end of a noun is associated with the lack of a phoneme "s" at the end of a following verb and conversely. Such associations hold for simple declarative sentences like:

5. The boy eats the rabbits.
6. The boys eat the rabbits.

Braine believes that the formation of associations, augmented by position learning, is adequate to explain how the syntax of simple declarative sentences is learned. In effect, Braine considers such sentences to be sequences of "primary phrases." Primary phrases are themselves sequences of open- and closed-class morphemes connected by associative bonds. "The location learned is that of a unit within the next-larger containing unit of a hierarchy of units. There are hierarchies at two levels: within sentences the units are primary phrases and sequences of primary phrases; within primary phrases the ultimate units are morphemes [p. 348]."

Braine is aware that the information that certain sequences of linguistic elements behave as units and that such units can appear only in specified positions in simple sentences does not exhaust the speaker's knowledge of syntax. There are many different kinds of sentences allowing nearly all possible orders of words and primary phrases. Thus, if we take into account *all* the types of sentences in which it may occur, there are indefinitely many permissible locations of a linguistic unit (see examples, sen-

tences 35-41). A list of the positions available to a linguistic unit could at best specify its behavior in only a circumscribed part of the language. Yet it is only the learning of such a list that contextual generalization could explain.

Braine meets this objection by restricting the scope of his theory to the assimilation of the grammatical properties of simple declaratives. He maintains this restriction is not arbitrary since simple declaratives have psychological and linguistic characteristics which justify postulating special processes for their assimilation. Braine thinks simple declaratives may predominate in the child's verbal environment, thus forming the primary models from which the child's knowledge of his language is extrapolated. Second, Braine claims recent work in linguistics divides grammar into two parts.

According to Harris (1957) and Chomsky (1957), the grammar of a language can be hierarchized into an elementary part, called the "kernel" of the language, and a second part which consists of a set of transformational rules for deriving complex sentences from simple ones. The kernel grammar contains the definitions of the main parts of speech and describes rules for constructing simple declarative statements . . . [p. 340].

Thus, if we can explain the acquisition of simple declarative sentences, we have accounted for the basic component of the grammar. The remaining portion—the complex sentences produced by transformation—is to be described as a set of *sublanguages*, one sublanguage for each type of sentence (passive, relative, question, etc.). Rather than attempting to study English in its full complexity, Braine concludes ". . . it seems that it would be sound strategy to aim first at finding an explanation for the learning of the kernel of the language, i.e., for the learning of the structure of the simple declarative English sentence. This constitutes enough of a problem already [p. 342]."

Finally, Braine argues that perceptual learning, of which contextual generaliza-

tion is a special case, is a primitive process which does not demand much in the way of intellectual capacity of the learner. Contextual generalization would therefore satisfy at least one requirement on any process involved in first-language learning, namely, that it "not require intellectual capacities obviously beyond the reach of the 2-year-old [p. 326]."

If the learning of syntax is the generalization of the ordinal positions in which linguistic units appear, it is evident that the initial stage must consist in the perceptual isolation of such units. Braine claims that an argument for the feasibility of contextual generalization is that the boundaries of such units can be identified with certain specifiable properties of the speech signal.

Braine proposes two sorts of cues the child could use to identify these boundaries. One is "intonation": The stress, rhythm, and pitch patterns of sentences are assumed to be acoustic features which communicate information about segmentation. The other is the position of closed-class morphemes which, Braine holds, tends to delimit phrases.

We now turn to a discussion of these arguments. We first consider the claim that simple declaratives ought to receive special treatment. Second, we investigate whether the linguistic character of simple declaratives can be selected by reference to the syntactic properties of sentential positions. Third, we ask how much of the relation between simple declaratives and other types of sentences can be expressed by such a specification. Fourth, we consider broader issues raised by Braine's treatment of inflection, intonation, and perceptual isolation of units. Finally, we discuss his experimental techniques and results.

THE ROLE OF DECLARATIVE SENTENCES

Because he believes that a theory of the simple declarative sentences explains the basic part of the grammar and that such sentences predominate in the child's linguistic environment, Braine holds an account of the learning of simple declara-

tives is important even if it does nothing else. We shall return to the question of the kernel grammar presently. Let us first consider the claim that the verbal environment of the child exhibits a preponderance of grammatical simple declaratives.

It is clear that normal speech among adults does not exhibit any statistical bias towards fully grammatical simple declarative sentences. On the contrary, adult speech is usually ungrammatical (cf. Maclay & Osgood, 1959), and there is little evidence that adults engage in a careful limitation of their linguistic output when conversing with children.³ Moreover, the verbal environment of children includes utterances produced by adults conversing among themselves, utterances produced by siblings with little command of the language, utterances heard on radio and television, etc. These diverse sources presumably form a heterogeneous verbal environment.

Even if simplified speech predominates in the child's verbal environment, there is no reason to suppose that the environment is unusually rich in simple declaratives. Analyses we have made of the speech of mothers taped during conversations with their children fail to support that hypothesis.⁴ On the contrary, of a

³ Brown and Bellugi (1964) do find a relatively large proportion of fully grammatical utterances in their recordings of mothers' speech to children. They do *not*, however, find that simple declaratives are preponderant among such utterances. On the contrary, in the only sample of their data they present (a sample which they say is "rather representative"), only one of the six sentences produced by the mother is of the simple declarative type.

⁴ The data were supplied by Margaret Bullowa and her staff at the Massachusetts Mental Health Center. The total represents 38 half hours of recorded conversation between three mothers and their children taped at ages ranging from 6 to 30 months. Six transcripts were selected for analysis, greater weight being given the recordings made at 20 months than those made at 6 months.

Neither our judgments of grammaticality nor sentential type were checked for inter-

total of 432 utterances, 258 were fully grammatical. Of these, only 46 were simple declaratives.

Of course the character of the verbal environment plays a major role in language acquisition. It determines which language, vocabulary, style, and accent the child learns. *What is unknown, however, is which features of the verbal environment are critical for such learning.* There is, at present, no reason to believe that the learning of English is facilitated by a preponderance of simple declaratives in the child's sample of his language. Nor is there any reason to suppose that such a bias normally obtains.

We turn now to the question of whether the simple declarative has any formal or linguistic peculiarities to which its claim for special psycholinguistic status might be referred.

Braine makes a mistake that has unfortunately been common in psychological investigations concerned with generative grammar.⁵ He supposes there exists a base or kernel grammar producing all and only simple declaratives and that the transformational operations producing complex sentences are defined over the declaratives generated by this base component. If this *were* the case, one

judge reliability. While some degree of latitude may be involved in judgments of the former kind, the criteria for the latter are reasonably objective. It is clear that there is need for an extensive survey of the verbal environment of the child; the data we have cited are intended only as preliminary.

⁵ For examples of discussions in which this mistake appears to have been made, see Miller (1962); Miller, Galanter, and Pribram (1960); Osgood (1963); and Mehler (1963). All these assume that linguistics assigns a privileged status to the simple declarative. For example, Miller's (1962) discovery that it takes less time to find the passive corresponding to an active than to find the passive corresponding to a question is *not* explained by appeal to the linguistic fact that the active is the underlying form in the production of the passive and the question. There is no such linguistic fact.

might plausibly maintain that the status of the simple declarative as the underlying *linguistic* form justifies a parallel psycholinguistic precedence.

Braine is correct in asserting that there is a base form from which all syntactically related sentences are directly or indirectly derived. It is also true that base form is produced by a set of rules whose formal properties distinguish them from other rules in a generative grammar. *But it is not true that the base form is the simple declarative sentence.* The kernel grammar does *not* produce simple declarative sentences; it does not produce *any* sentences. Rather, the kernel grammar produces abstract structures that are transformed into a variety of different sentence types of which the simple declarative is one. In particular, the kernel sentence discussed by Chomsky (1957) should not be confused with these abstract structures. Kernel sentences differ from sentences of other types solely in that they are the consequence of applying only obligatory transformations to the kernel structure. Kernel sentences are thus in no sense the source for, or underlying form of, sentences of other syntactic types.⁶

Since the misunderstanding of the kernel notion has been widespread, it is worth indicating some of the linguistic considerations that militate against supposing the simple declarative to be the underlying form from which other sorts of sentences are derived. Consider the passive construction. We might attempt to derive 7*b* from 7*a* by a rule like that given in 8.

- 7*a*. The boy chases the dog.
7*b*. The dog is chased by the boy.

⁶ That this has always been Chomsky's view is clear from a reading of *Syntactic Structures*. That Harris does not hold the simple declarative to be the base form follows from the fact that the notion of a derivation plays no role in Harris' theory. The mappings Harris (1957) employs in transformational analysis are characteristically symmetrical, hence there can be no questions of identifying an underlying syntactic form.

8. If NP_1 Verb NP_2 is a declarative sentence, and if NP_2 is the object of the verb, then NP_2 is Verbed by NP_1 is the corresponding passive.⁷

But now, consider 10, the result of applying 8 to 9:

9. The boy chases the dogs.

10. *The dogs is chased by the boy.

a string which is not grammatical for many dialects since the number of the verb should be determined by the subject. To avoid 10, Rule 8 must be split into two rules:

8a. NP_1 V NP_2 + sg \rightarrow NP_2 is V + ed by NP_1 .

8b. NP_1 V NP_2 + pl \rightarrow NP_2 + s are V + ed by NP_1 .

However, consider 12 and 14, the result of the application of 8a to 11 and 13 respectively.⁸

11. The boy is chasing the dog.

12. *The dog is is chasing ed by the boy.

⁷We shall adhere to the notational conventions employed by linguists. In particular, ungrammatical strings will be preceded by *. \emptyset stands for the zero number of a linguistic class (the plural of the English word "sheep" is thus "sheep + \emptyset "). The following abbreviations will also be employed: NP for Noun Phrase, T for Article, VP for Verb Phrase, S for Sentence, N for Noun, V for Verb, sg for the singular morpheme, pres for the present-tense morpheme, pl for the plural morpheme, Det for determiner. Be will be used to designate any inflection of the verb "to be."

⁸If 8a can be allowed to interpret *is chasing* as a V at all, then it produces the incorrect form *12; if it cannot, then another new rule is required to produce 19 from 11. In all the examples in this paper, we do not claim to present the unique solutions and rules, either for those formulations which we show to be essentially incorrect or for those that are essentially correct. All the rules are considered out of the context of a presumed full grammar. In that context they might appear somewhat differently—but the distinctions and characteristics with which we are concerned will remain unchanged.

13. The boy chased the dog.

14. *The dog is chased ed by the boy.

Just as the difference in the number of the object required us to adopt different passive rules for 7a and 9, so two more rules, 8c and d, will be required to passivize 11 and 15:

15. The boy is chasing the dogs.

8c. NP_1 is V ing NP_2 + sg \rightarrow NP_2 + sg is being V ed by NP_1 .

8d. NP_1 is V ing NP_2 + pl \rightarrow NP_2 + pl are being V ed by NP_1 .

and similarly for 13 and 16.

16. The boy chased the dogs.

8e. NP_1 V + ed NP_2 + sg \rightarrow NP_2 + sg was V + ed by NP_1 .

8f. NP_1 V + ed NP_2 + pl \rightarrow NP_2 + pl were V + ed by NP_2 .

In general, if we derive passives from their corresponding actives, a different passive rule is required for each choice of object number and verb tense. For five tenses and two numbers there are at least 10 rules required to derive the simple passive from declarative sentences. Furthermore, even these 10 rules will not serve to derive the passive of more complicated sentences. For example, we will need special rules for:

16a. Does the boy chase the dog?

16b. Is the boy being chased by the dog?

17a. Why does the boy chase the dog?

17b. Why is the dog chased by the boy?

and so on.

In each of the cases we have discussed, the problem clearly arises from the attempt to derive the passive from its corresponding declarative. This difficulty would be avoided were it possible to define the transformation which rearranges the subject and object phrases so that it applies prior to the attachment of tense and number to the verb. In this way we specify the operations on the noun and verb relevant to passivization *without reference to the particular*

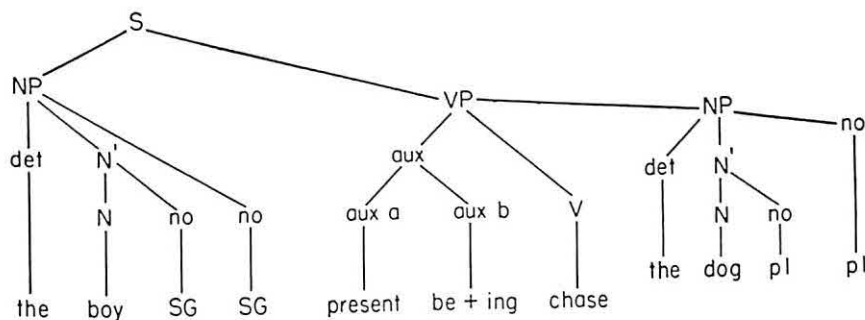


FIG. 1. Tree diagram of the underlying (untransformed) structure of the sentence:
"The boy is chasing the dog."

choice of tense and number they exhibit in a given sentence. This rule is easily formulated:

18. NP_1 aux V NP_2 is rewritten as NP_2 aux be + Past part. V by NP_1 .

However, it can be applied to generate all passives only if suitable abstract representations of sentences are provided as its domain. For example, the passive sentence, 19:

19. The dogs are being chased by the boy.

has, as its corresponding active, Sentence 15. Yet, if 19 is to be generated by an application of Rule 18, 15 cannot be its source, since 15 does not possess the formal properties required of strings in the domain of 18. No definite sense can therefore be given to the notion of applying 18 to 15. Rather, the source for 19 is the abstract illustrated in Figure 1.⁹

⁹ Of course, the result of 18's application to the structure in Figure 1 is *not* an actual sentence. The full derivation of Sentence 19 from the structure represented in Figure 1 is accomplished in at least three steps: (a) Rule 18 converts it into the string (constituent structure is not marked here) *the dog pl pres be+ing be+Past part. chase by the boy sg sg*; (b) an affix-attachment rule (affix, $V \rightarrow V + \text{affix}$) then applies to produce the string *the dog pl be+pl+pres be+ing chase+Past part. by the boy sg sg*; (c) convert morphophonemic "spelling" rules convert this into Sentence 19 (see the paragraph following 33 in this paper and Chomsky, 1957, for other examples of affix movement and sentence derivations). The concept of the

It can be demonstrated that derivations employing Rule 18 can account for the examples we have investigated so far. This means, in effect, that we have eliminated the proliferation of passive rules by insuring that inflections of number and tense are not specified by the passive rule. To do this, however, we have had to assume that the underlying form from which the passive derives is not a corresponding active but rather an abstract structure never realized in speech. Since analogous considerations show that the active itself is merely one of the deformations of this underlying structure, there would appear to be nothing in the linguistic theory of the derivation of English sentences which would justify assigning a special status to active declarative sentences.

abstract and concrete level in sentences is a formal representation of many intuitions: for example, the difference between "logical" and "apparent" grammatical relations, the description of relations among sentences, and so on (see Fodor & Katz, 1964, for various discussions). Notice that by having the affix-attachment rule follow the passivization rule, 18, we avoid formally the multiplication of passive rules. Thus, this analysis is a formal account of the intuition that "passive" sentence is a unitary notion, and the intuition that it is the number of the apparent subject of the verb which determines verb number in English, not the number of the underlying or "logical" subject (see the section on inflection and Footnote 10 in this paper).

CONTEXTUAL GENERALIZATION AND THE LEARNING OF POSITION

Suppose, however, that we grant Braine's assumption that there is some point to a special theory of the psychological processes underlying the learning of the syntax of declaratives. We must ask how much of the character of such sentences emerges from an analysis of the order relations among their constituents. We shall see that very little of the speaker's knowledge of the syntactic structure of simple declaratives can be attributed to having learned such relations.

To see how little of the syntax of a simple declarative is expressed by order, it is only necessary to consider sentences whose grammatical structure is not so expressed. Thus a speaker who knows that 20 and 21 are sentences also knows that while 22 is perfectly good English, there is something wrong with the syntax of 23.

20. The kangaroo cost 10 dollars.
21. The child blew the kazoo.
22. The kazoo was blown by the child.
23. *Ten dollars was cost by the kangaroo.

This example is particularly embarrassing for Braine, since 20 and 23 exhibit precisely the sort of positional relations on the basis of which contextual generalization is supposed to operate. That is, since both "10 dollars" and "the kazoo" appear postverbally, and since "The kazoo" appears preverbally in 22, the operation of contextual generalization ought to permit the preverbal appearance of "Ten dollars" in 23. The ungrammaticality of 23 suggests that there must be a difference between "blow" and "cost." Though this difference is not revealed by their positional privileges in simple declaratives, it is decisive for determining how passivization operates in sentences in which they occur.

A similar case is the following:

24. John phoned Jane up.
25. John phoned up Jane.

26. John phoned her up.
27. *John phoned up her.

Since "up" appears in the fourth and third positions respectively in 24 and 25, and since it appears in the fourth position in 26, Braine's theory predicts that place generalization requires its occurrence in 27. This prediction is incorrect. On the contrary, it is a necessary condition upon the grammaticality of sentences of the form:

28. John phoned up X.

that X not be a pronoun. Hence, if we are able to formulate the rule which permits 25 but precludes 27, we must take account not only of the ordinal positions of the words in those strings, *but also of the syntactic classes to which the words belong.*

These examples are a consequence of a quite general fact: *the types of expression that can appear at a given ordinal position in a simple declarative sentence are extremely heterogeneous.* This fact is of the utmost importance for an evaluation of Braine's theory. For it entails that contextual generalization is an inadequate mechanism for extrapolating the grammatical regularities the child observes in his language. Many types of expression can appear at a given sentential position. Hence sentences exhibiting precisely the same ordinal relations among their segments may nevertheless be of different syntactic types. Since the operation of contextual generalization requires only that features of order be common to conditioned stimulus and generalized stimulus, it follows that we cannot infer that String S is a sentence whenever there exists a sentence, S', related to S in a way that satisfies the conditions for contextual generalization.

To put it slightly differently, the examples just discussed demonstrate that some of our information about the syntactic character of simple declaratives cannot be expressed in terms of location information; rules which determine the interrelations among simple declaratives and the relations between simple declaratives and other sorts of sentences

distinguish between expressions that can appear in a given location.

The similarity between Braine's view and that of Jenkins and Palermo (1964) is most evident at this point. Jenkins and Palermo apparently hold that the syntactic structure of a sentence can be expressed by a sequence of markers representing the classes of which the words comprising the sentence are members. Thus, they remark that "... the ideas we want to present here are of two sorts: *sequence* and *class* [p. 164]" and that

"Colorless green ideas sleep furiously" is a sentence in English not because it is true, or sensible, or interpretable by the listener, but because it is a "correct" assembly of classes appropriately modulated (*i.e.*, they are the right general classes of entries properly modified to take their places in the particular sequence). . . . The critical question, then, is seen to be that of the organization of the elements into classes [p. 164].

It is thus evident that Jenkins and Palermo's view of syntax is simply a weaker version of Braine's. For while Jenkins and Palermo acknowledge only sequences of class markers (and inflection), Braine has noticed the necessity of providing some psychological mechanism to account for the phrase structure of such sequences.

In short, there is no basis for Jenkins and Palermo's belief that the employment of mediational paradigms affords a breakthrough in the study of syntax assimilation. Though the formation of word classes can perhaps be accounted for by such paradigms, they throw no light whatever upon the assimilation of even such relatively superficial syntactic features as phrase structures. It goes without saying that appeals to them are utterly unilluminating on the question of how the child learns the deep syntactic structures with which this paper is primarily concerned.

THE DERIVATION OF ORDER

Thus far we have noted one sense in which order is a relatively unimportant

feature of syntactic structure even in English where inflection is not widely used: Identity of order relations is compatible with considerable differences in syntactic form. There is a more important point which also stems from Braine's failure to distinguish between underlying and surface structure. Often a correct formulation of the rules determining syntactic structure requires distinguishing between the order of lexical items in the sentence and in the underlying representation from which the sentence is derived. While contextual generalization might conceivably account for learning the former, it is patent that it could not account for the learning of the latter.

For example, in our discussion of the passive we treated *be + ing* as a unit in the underlying representation of sentences like 15. There are intuitive and formal advantages to this treatment. It permits us to account for the intuition that *be + ing* functions as a semantic unit indicating a particular mode of the main verb. Moreover, treating these morphemes as a single item permits us to account for the fact that such strings as:

29a. *The man is chase the dog.

29b. *The man chasing the dog.

are not sentences: If *be + ing* is represented as a lexical unit, one of the morphemes comprising that unit cannot be selected without also selecting the other.

Cases in which simplicity of representation and linguistic intuition require an underlying order differing from the order of elements in the manifest sentence are found throughout English. Thus in sentences 24-27, the underlying form of the verb must be *phone + up*, since English restricts the particles that can accompany certain verbs. Thus we have no

30a. *John phoned down the girl.

30b. *John phoned the girl down.

Similarly, 31a, but not 31b, is grammatical.

31a. They looked the house over.

31b. *They looked the house in.

To account for such examples and for the fact that the verb-particle sequence forms a semantic unit, such sequences are recorded in the underlying form as *verb + particle*. But notice that, though the particle precedes the object in the *underlying* representation of 27, a mandatory transformation of that underlying structure permutes the particle and the object whenever the latter is a pronoun. Hence the underlying form of 27 has a different order from its manifest form. Moreover, this underlying order is *never* directly reflected in an actual sentence.

Finally, even the very simplest sentences derive from an underlying form in which the order of the elements differs from the manifest order. Consider:

32. The child runs.

The terminal portion of the tree representing the underlying structure of this sentence is:

33. Det N sg pres V.
Øə/čaild/run.

The tense marking (present) precedes the verb phrase in the underlying structure in order to permit a uniform treatment with more complicated sentences. In the manifest sentence, the tense marker is always attached to the first verbal element in the verb phrase whether that element is an auxiliary, modal, or compound verb. This is accomplished for complicated expansions of the auxiliary by a rule which permutes the tense marker with the adjacent verbal element to the right. If the tense and verb order for 32 were not as represented in 33, 32 would constitute an exception to this rule and would thus require a complication of the grammar (cf. Chomsky, 1957, and Footnote 9).

This indirect relation between the order of elements in the underlying form and their order in the manifest sentence poses serious problems for *any* theory of syntax learning. It is clear that the child never encounters manifest models

of the underlying sentence order. Adults do not utter such sequences as:

34. Øə čaild s run.

We have seen, however, that in a formalization of English grammar 34 represents a step in the derivation of 32. Hence, if the child who learns English learns the rules governing the syntax, it follows that he *must learn to manipulate underlying structures, for which his verbal environment provides him with no explicit models*. It is evident that contextual generalization, simply because it is a variety of generalization, cannot account for such learning.

"SUBLANGUAGES"

We have seen that relatively little of the syntactic character of simple declarative sentences can be expressed in terms of the manifest order of their constituents. We now ask how much of the syntactic relations between simple declarative and other types of sentences can be expressed in those terms. We maintain that the behavior of a linguistic unit in complex sentences cannot be captured by a theory which represents its syntactic properties by a list of the positions it is capable of occupying. There are two reasons for this, one of which Braine acknowledges; the other has escaped him.

It is clearly possible to construct complicated sentences exhibiting almost any order of syntactic constituents. Thus in Sentence 15, Braine would presumably distinguish three major positions occupied, respectively, by *the boy*, *is chasing*, and *the dogs*. The relations between these three positions are however, very flexible. Consider:

35. The dogs are being chased by the boy.

36. It is the dogs the boy is chasing.

37. What is being chased by the boy are the dogs.

38. What the boy is chasing are the dogs.

39. It is the dogs that are being chased by the boy.

40. Chasing them is what the boy is doing to the dogs.

41. Chasing the dogs is what the boy is doing.

Far from it being the case, as Braine suggests, that transformation tends to produce relatively minor variations in order, examples 35-41 show that order is the syntactic property transformation is least likely to preserve.

Braine is aware that relative order is preserved only within sentences of the same syntactic type. This leads him to suggest that natural languages should be thought of as systems of "sublanguages," distinguished by a characteristic constituent order. This suggestion is defective in a number of ways. First, if sublanguages are specified solely by the ordinal relations of their constituents, there must be an infinity of sublanguages since sentences in a natural language may be arbitrarily long, and every pair of sentences which differ in length will ipso facto differ in constituent order at some level. It is thus logically impossible that the child should master his language by learning one sublanguage after another.

The preceding demonstrates what was shown above on purely syntactic grounds: The differences between types of sentences cannot, in general, be specified in terms of differences in constituent orders since sublanguages cannot be defined in those terms.

This can be seen most strikingly in the ambiguity of such sentences as:

42. The office of the president is vacant.

where the lexical items and order of constituents *at all levels* are the same on both readings despite the fact that any reasonable analysis of English into sublanguages must assign 42 to two sublanguages, because it can be questioned in two quite different ways (cf. 43 and 44):

43. Where is the office of the president?

44. Who held the office of the president?

There are other examples which illustrate the impossibility of exploiting constituent order to capture relations between syntactically different types of sentences: 45 is related to 46 rather than to 47 while the reverse holds for 48.

45. John is easy to please.

46. They please John.

47. John pleases them.

48. John is eager to please.

Since, however, the constituents of 45 and 47 are ordered in precisely the same way, we cannot appeal to order to explain the relations between the sublanguages these sentences belong to.

These examples reflect a fact that must be evident to anyone who has seriously considered the problem of describing the syntactic structure of a language like English: Very few of the interrelations between types of sentences are expressed by similarities and differences of manifest order. Rather, a theory which marks such interrelations must do so in terms of highly abstract constructs such as the phrase analyses and transformational histories underlying sentences. Linguistic evidence shows that no simpler apparatus explains intersentential relations that are intuitively evident to native speakers. It follows that Braine's attempt to specify sublanguages and their interrelations in terms of constituent order can be of no serious explanatory value.

INFLECTION

We now turn to some of the broader issues raised by Braine's theory. The first arises from difficulties implicit in Braine's assumption that the processes employed in learning order languages are different in kind from those employed in learning inflected languages.

The search for the mechanism of syntax acquisition is, presumably, a search for species-specific capacities common to all children. An adequate theory of syntax acquisition must explain the ability of a child to learn any language to which he happens to be exposed. It follows that a theory of language learning which makes essential reference to

intonation is intimately related to syntactic structure, this indicates that the order of assimilation of syntactic structures is unlikely to be the one mentioned above. Instead, it suggests that derived constituent structure may be among the earliest syntactic information assimilated.

In short, the correct interpretation may be not that the perceived location of pause, stress, and intonation are the child's clue to the analysis of structure, but rather that the prior analysis of structure is what determines where the child learns to hear pause, stress, and intonation. Nor is the possibility of some intermediate position excluded.

EXPERIMENTATION ON CONTEXTUAL GENERALIZATION

We have considered some inadequacies of explanations of syntax assimilation that appeal to contextual generalization. However, Braine has shown that subjects *can* learn miniature artificial languages whose constituents are nonsense words by position generalization. This may appear to provide a foundation for theories of language learning in which contextual generalization plays an important role. However, Braine's results are equivocal. A brief discussion of his experimental techniques is therefore in order.

There is some danger in applying results garnered in the study of the learning of artificial languages by latency-aged children to the learning of first languages by preschoolers. First, it is possible that the psychological processes mediating the learning of first languages are pretuned for the assimilation of systems having quite specific formal properties. If this is correct, very little will be revealed about these processes by studying the assimilation of artificial languages whose structure is arbitrarily different from that of natural languages.

Second, as Thorpe (1961) has suggested, the learning of language may be one of the capacities for which

there . . . exist specific brain mechanisms ready to be activated during and only during a particular period of the life span of the child and . . . if they are not properly activated at the right time subsequent activation is difficult or impossible, resulting in permanent disabilities in later life [p. 200; cf. also Penfield & Roberts, 1959].

This speculation supports the widely held view that adults learn second languages in a manner essentially different from the way children learn first languages.

Braine has attempted to control for this source of error by replicating the simpler of his experiments with a population of 4-year-olds. The point is not, however, that very young children may be incapable of contextual generalization. It is rather that there may be language-learning processes operative in younger children which do not occur in older ones. Braine controls for the former but not the latter of these possibilities.

We now consider the details of Braine's experiments. It is essential to Braine's argument to demonstrate the existence of contextual generalization among phrases as well as words. We have seen that Braine defines the notion of a position in such a way that it may be occupied by a word in one sentence and by a phrase in another.

Braine's experiments are thus divided into two groups. The first demonstrates contextual generalization in a language whose sentences are all of the type shown

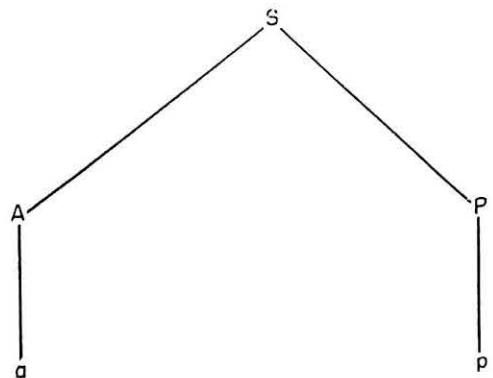


FIG. 2. Tree diagram of the alleged structure of nonsense sentences in Braine's *a + p* language.

(around 1-1½ years) as the period when the child "speaks without words."

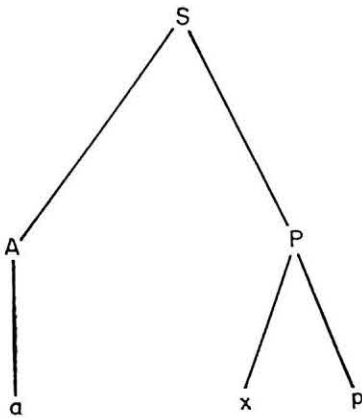


FIG. 3. Tree diagram of the alleged structure of the nonsense sentences in Braine's $a + xp$ language.

in Figure 2, where A and P are classes of single (nonsense) words, and a and p are items drawn from those classes.

Braine's second experiment purports to show that contextual generalization can also function when the constituents are phrases. Here Braine uses language in which sentences are said to be either of the type shown in Figures 3 or 4.

But Braine provides no support for the analyses of the String axp in Figures 3 and 4 beyond the remark that, in the situations where ax was intended to have phrase status, a and x were presented together, while in the situation where xp was intended to have phrase status, x and p were presented together.

In short, Braine's experimentation is at least compatible with the interpretation that what his subjects *learned* were strings with the structure exhibited in Figure 5. So long as this interpreta-

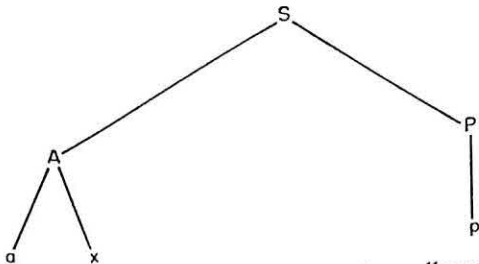


FIG. 4. Tree diagram of the alleged structure of the nonsense sentences in Braine's $ax + p$ language.

tion remains open, Braine can hardly claim that his experiments demonstrate contextual generalization among phrase-length constituents.

In experiments with artificial languages consisting of a finite set of strings, it is extremely difficult to support claims about the psychological reality of abstract structures like phrases. In natural languages, appeals to the intuitions of speakers concerning the appropriate segmentation of sentences constrain phrase analyses to some extent. More important, formulation of the rules for transformation and stress imposes certain requirements upon the phrase-structure analysis since the phrase structure articulates the domain to which these rules apply. In the case of the artificial languages Braine studies, neither of these considerations holds. Speakers do not have structural intuitions

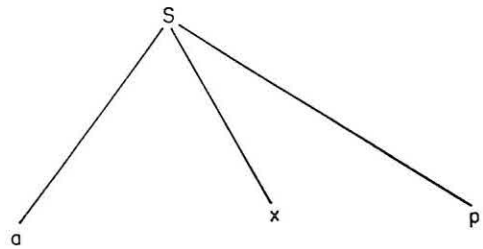


FIG. 5. A possible alternative structure for sentences allegedly belonging either to $a + xp$ or to $ax + p$ languages.

about artificial language, and no language which consists of a finite set of strings requires phrase-structure rules in its grammar, for any such language can be enumerated by a simple list. If, therefore, it is suggested that Braine's subjects may not have analyzed the stimulus material in the way Braine wanted them to, it is difficult to see what reply he could make.

The difficulty may have stemmed from Braine's belief that the process responsible for the perceptual integrity of the phrase is association. On that view, co-presentation of pairs of nonsense syllables would perhaps tend to establish such pairs as phrases. The fact that the pairs of items intended to be perceived

as phrases were presented together would thus help to justify the analysis provided in Figures 3 and 4.

The claim that the integrity of the phrase is mediated by processes of association is, however, totally untenable. Phrases in natural language can exhibit grammatical dependencies across strings of arbitrary length. It is thus inconceivable that the psychological unity of phrases could be attributed to associative bonds between the elements. Conversely, the suggestion that the basis of phrase integrity is associationistic in short phrases but that some different and unknown psychological mechanism operates in longer phrases is too evidently ad hoc to merit considerations.

REFERENCES

BRAINE, M. D. S. On learning the grammatical order of words. *Psychological Review*, 1963, 70, 323-348.

BROWN, R., & BELLUGI, U. Three processes in the child's acquisition of the syntax. *Harvard Educational Review*, 1964, 34(2), 133-151.

CHOMSKY, N. *Syntactic structures*. The Hague: Mouton, 1957.

CHOMSKY, N., & MILLER, G. Introduction to the formal analysis of natural languages. In R. Luce, R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology*. Vol. 2. New York: Wiley, 1963. Pp. 269-323.

FODOR, J., & KATZ, J. (Eds.) The structure of language. In *Readings in the philosophy of language*. Englewood Cliffs, N. J.: Prentice-Hall, 1964. Pp. 479-518.

GRÉGOIRE, A. *L'Apprentissage du langage*. Liège, Belgium: l'Université de Liège, 1937.

HARRIS, Z. Co-occurrence and transformation in linguistic structure. *Language*, 1957, 33(3), 283-340.

JAKOBSON, R. *Kindersprache, Aphasie und allgemeine Lautgesetze*. Uppsala, Sweden: Uppsala Universität, 1941.

JENKINS, J., & PALERMO, D. Mediation processes and the acquisition of linguistic structure. In Ursula Bellugi & R. W. Brown (Eds.), *The acquisition of language*. Chicago: Univer. Chicago Press, 1964. Pp. 141-169.

LEES, R. The grammar of English nominalizations. *International Journal of American Linguistics*, 1960, 26(Pt. 2), 205-220.

LENNEBERG, E. The capacity for language acquisition. In J. Fodor & J. Katz (Eds.), *Readings in the philosophy of language*. Englewood Cliffs, N. J.: Prentice-Hall, 1964. Pp. 579-608.

LIEBERMAN, P. On the acoustic basis of the perception of intonation by linguists. *Word*, 1965, 21, in press.

MACLAY, H., & OSGOOD, C. Hesitation phenomena in spontaneous English speech. *Word*, 1959, 15, 19-44.

MCCARTHY, D. Language development in children. In L. Carmichael (Ed.), *A manual of child psychology*. (2nd ed.) New York: Wiley, 1954. Pp. 429-630.

MEHLER, J. Some effects of grammatical transformations on the recall of English sentences. *Journal of Verbal Learning and Verbal Behavior*, 1963, 2, 346-351.

MILLER, G. A. Some psychological studies of grammar. *American Psychologist*, 1962, 17, 748-762.

MILLER, G. A., GALANTER, E., & PRIBRAM, K. H. *Plans and the structure of behavior*. New York: Holt, Rinehart & Winston, 1960.

OSGOOD, C. E. On understanding and creating sentences. *American Psychologist*, 1963, 18, 735-751.

PENFIELD, W., & ROBERTS, L. *Speech and brain mechanisms*. Princeton: Princeton Univer. Press, 1959.

THORPE, W. H. Sensitive periods in the learning of animals and men. In W. H. Thorpe & O. L. Zangwill (Eds.), *Current problems in animal behavior*. Cambridge: Cambridge Univer. Press, 1961. Pp. 194-224.

(Received August 31, 1964)

ON THE BASIS OF PHRASE STRUCTURE: A REPLY TO BEVER, FODOR, AND WEKSEL

MARTIN D. S. BRAINE

Walter Reed Army Institute of Research, Washington, D. C.

On Bever et al.'s central contention, the linguistic evidence seems insufficient to show that the order of elements in the simple declarative sentence is a deformation of the order in the underlying string: The deformation appears to be imposed by the choice of a particular phrase-structure model, rather than a fact about English. While undoubtedly insufficient and imprecise, the writer's proposals on grammar acquisition seem to have more potentialities than Bever et al. allow. Too little is known for it to be argued that they are inconsistent with the character of the verbal environment. Limitations of the proposals in respect to the learning of word classes and several other issues raised by Bever et al. are discussed.

The arguments of Bever, Fodor, and Weksel (1965b) will be discussed in the order in which they were presented.

THE VERBAL ENVIRONMENT

Bever et al. begin by saying that I (Braine, 1963) claim "the verbal environment of the child exhibits a preponderance of grammatical simple declaratives [p. 470]," and they cite data to the contrary. Their data also indicate that the child is exposed to many ungrammatical utterances, although there appears to be some disagreement with Brown and Belugi (1964) on this point.

I did not claim that children were mainly exposed to simple declaratives, and it greatly distorts my conception to think that a simple count of the number of simple declaratives in adult speech is relevant to it. Information about positional and contingency relations in simple sentences is abundantly exemplified in many transforms (as well as in many "not-fully-grammatical" utterances). As one among many examples that could be cited, consider the similarity in verb-phrase structure between passives and sentences with adjective predicators (e.g., *George was served—George was sensible*): The same auxiliary structure (*has been, would be, might have been being*, etc.) occurs with past participles and with adjectives. Experience with the verb structure of one sentence type could hardly fail to trans-

fer to the other type. In general, it seems to me that too little is known about the verbal environment and the child's response to it to build much of an argument for or against my proposals.

The conflicting data on the prevalence of ungrammatical utterances raise some question as to the criteria of grammaticality used. It seems important to distinguish at least three kinds of not-fully-grammatical utterances: (a) incomplete sentences where the speaker hesitates or changes his mind in midsentence—these are likely to be grammatical between hesitations, (b) correctly constructed phrases (isolated noun phrases, prepositional phrases, etc.) which expose the major components of sentences to the learner uncluttered by context and may be important in the learning of the segmentation of sentences, as well as in the learning of the internal structure of the phrases themselves, (c) radically ungrammatical utterances with the appearance of being random combinations of words (e.g., *To went school he* for *He went to school*). It is primarily this last category, which seems very infrequent, that might be expected to handicap learning.

MANIFEST AND UNDERLYING STRUCTURE

The next argument is central to the critique. It has two parts: Bever et al. claim I made a mistake in supposing (a) that there is "a base or kernel grammar

producing all and only simple declaratives," and (b) that "the transformational operations producing complex sentences are defined over the declaratives generated by this base component [p. 471]." It should be carefully noted that (b) is only very tangentially relevant to my article. My discussion of transformations was too brief to be more than approximate, and its main purpose was to show that the existence of transformations did not contradict my proposals concerning simple declaratives. Yet most of the arguments of Bever et al. relate to (b) rather than (a). In particular, their long discussion of the passive transformation, associated with sentences 7a-19, is relevant only to the question of what the passive transformation is a transformation of. Bever et al. argue that certain rules (e.g., the subject-verb concord) apply to both active and passive, so that much duplication would be avoided in a grammar if these rules applied *after* the transformation was made. This argument says nothing about the relation between the kernel grammar and simple declarative sentences. Bever et al. implicitly admit its irrelevance when they say "analogous considerations show that the active itself is merely one of the deformations of this underlying structure [p. 473]." Whether the simple declarative is a "deformation" of an underlying terminal string thus turns on what these unstated "analogous considerations" are. It is difficult to take issue with unstated arguments. The discussion below relates to English, for which I am fairly clear about what the analogous considerations are.

I know of only one place in the generative grammar literature where a direct argument is made that transformations are involved in the generation of the simple declarative English sentence—that is Chomsky's (1957, pp. 38-42) argument relating to the auxiliary verb structure. Chomsky makes it clear that the validity of his argument turns on the question of whether discontinuous constituents are permissible in a phrase-structure grammar. He concludes: "In the auxiliary verb phrase we really have discontinuous elements, e.g., . . . *have* . . . *en* and *be*

. . . *ing*. But discontinuities cannot be handled within [phrase-structure] grammars [p. 41]." Whether or not a phrase-structure (kernel) grammar can permit discontinuities is a technical question about the form of such grammars. Since Bever et al. rest so much of their case on the argument that the simple declarative sentence does not reflect the order of elements in the underlying string, it is necessary to examine their conception of a phrase-structure grammar in detail.

Phrase-structure grammars employ symbols representing morpheme and phrase classes (conventionally written in capital letters) and symbols representing morphemes (conventionally written lower-case). Rules are of the form $\varphi \rightarrow \psi$ (to be read " φ may be rewritten as ψ "), where φ and ψ are strings of symbols. There is a privileged symbol, S (= "sentence"), and the first rewriting rule applied in generating a sentence is of the form $S \rightarrow \dots$; each rewriting rule permits one capital letter to be rewritten, and the rewriting rules are successively applied until all capital letters have been rewritten, leaving a string of morphemes which is the sentence or "terminal string" generated. The grammars defined in this manner comprise a large family of possible models. Different types of models in the family are specified by defining kinds of rewriting rules that are permitted (see Chomsky, 1963, for a description of several types). The permissibility of discontinuous rules turns on whether it is required in expansion rules $A \rightarrow \psi$ that ψ be a continuous string. Discontinuous rules drop this requirement to the extent of permitting such expansions as $A \rightarrow \lambda \dots \mu$, where the ellipsis indicates that λ and μ straddle the element on the right of A. Thus, a pair of rules, $X \rightarrow AY$, $A \rightarrow \lambda \dots \mu$, generates strings of the form $\lambda Y \mu$ in which the components are Y and $\lambda \dots \mu$. A grammar which forbids discontinuous rules is obliged to generate such a string in two stages: Phrase-structure rules, $X \rightarrow AY$ and $A \rightarrow \lambda \mu$, generate terminal strings with the order $\lambda \mu Y$, and a transform then permutes the order $\lambda \mu Y$ into $\lambda Y \mu$. Such a grammar is forced to impose on the language

a distinction between manifest structure ($\lambda Y\mu$) and "underlying structure" ($\lambda\mu Y$). It is now known that certain kinds of discontinuities, which include those that appear in simple English sentences, can be tolerated in phrase-structure grammars (Matthews, 1963; Yngve, 1960).¹ The recent work makes it clear that Chomsky's (1957, p. 41) statement that "discontinuous rules cannot be handled within [phrase-structure] grammars" is wrong. The use of models which do not admit discontinuities is controversial within linguistics.² Discontinuous constituents have been widely used by grammarians outside the generative grammar school (e.g., Bloch, 1946; Harris, 1951; Wells, 1947).

In their section, "The Derivation of Order," (which continues the argument begun in their earlier section, "The Role of Declarative Sentences") Bever et al. cite as evidence for the distinction between manifest and underlying order the phrases *be chasing* and *phone her up*.

¹Yngve's discontinuous rules are of the form $A \rightarrow \lambda \dots \mu$, with the restriction that sentences are to be generated from left to right. Matthews defines a much wider class of discontinuities. The requirement of left-to-right generation could probably be dropped, although specification of an order of application of rules is necessary. (In the above example, if $Y \rightarrow WZ$, then the rule sequence $X \rightarrow AY$, $A \rightarrow \lambda \dots \mu$, $Y \rightarrow WZ$ generates $\lambda WZ\mu$, and the rule sequence $X \rightarrow AY$, $Y \rightarrow WZ$, $A \rightarrow \lambda \dots \mu$ generates $\lambda W\mu Z$: Failure to specify an order of application would leave the terminal order vague.) Specification of an order among rules is a normal practice in grammars without discontinuities (Bach, 1964). Matthews shows that, with the left-to-right restriction, grammars with discontinuous rules generate context-free languages; without this restriction they generate context-sensitive languages.

²Unfortunately, the argument over discontinuous rules has taken place not over whether such rules exist and what their precise form is but rather as part of the larger argument that transformational grammars are necessary (e.g., Harman, 1963; Postal, 1964). The likelihood that both transformational and discontinuous rules exist has received little discussion.

Both these are of the $\lambda Y\mu$ form. A grammar which forbids representing the constituents in the form *be . . . ing* and *phone . . . up* must necessarily generate the orders *be+ing+chase* and *phone+up+her*; having generated the items in the wrong order, it is of course necessary to permute them into the actual order. The entire argument associated with sentences 29a-34 rests on a particular choice of phrase-structure model.

This distinction between manifest and underlying structure in kernel sentences creates a special kind of methodological difficulty in assessing the validity of generative grammars. A full grammar contains both a phrase structure and a transformational component. If the terminal strings generated by the phrase structure are permitted to be arbitrarily different from any actual sentence structures, there are no independent data against which the phrase structure and the transformational rules can be separately tested. Since transformational rules provide an extraordinarily powerful tool for mapping one system into another, the grammarian can write the phrase-structure kernel partly on the basis of his convenience, free to correct any poor fit with the manifest structure of the language by using transforms to reshuffle the elements. This methodological looseness makes it impossible to accept empirical claims about the properties of phrase-structure grammars of natural languages. (For example, the similarity in structure between ordering and inflected languages, reported by Bever et al. in their section, "Inflection," may not reflect a linguistic universal, but rather the prior decision to treat a certain kind of phrase-structure description as convenient.)

In general, in order to establish the existence of "deformations" in simple declaratives, Bever et al. would have to show that there are no feasible forms of phrase-structure model which would permit transformational grammars to be written so that only the identity transform intervened between the phrase structure and the morphophonemic rules. No one has yet attempted to make such a difficult argument. Instead, it seems to have been

assumed that, if transformational rules are to be used at all, then there is no motive for not exploiting them, leaving the detailed character of phrase-structure grammars to be settled, in part, according to the grammarian's convenience. It is doubtful, however, that the distinction between manifest and underlying structure in kernel sentences is ultimately convenient even to the grammarian, for the methodological reasons noted above. The distinction is certainly not convenient psychologically: The assumption that the learner learns structures to which he was never exposed would indeed pose "serious problems for *any* theory of syntax learning [p. 476]," as Bever et al. note.

THE SCOPE OF THE PROPOSALS

The next argument, associated with sentences 20–28, is that my proposals account for very little of the structure even of simple declarative sentences. Although my article did not claim sufficiency, my proposals seem to account for more of the structure of the simple declarative than Bever et al. allow.

The assertion that the theory must fail for sentences 24–28 reflects an over-simplistic interpretation of the theory. Positions learned are those determined by the immediate constituent analysis; I never remotely suggested they were a simple 1-2-3-4 ordinal series. *John phoned Jane up* contains two main parts—*John* and *phoned Jane up*; *phoned Jane up* contains two parts—*phoned . . . up* and *Jane*. Presumably English speakers learn that verb units often contain verb-particle sequences with the particle frequently contingent on the verb (*phone up* but not *phone down*, *think over* but not *think across*, etc.), and they also learn that noun phrases sometimes follow the verb-particle sequence (*phone up the girl*) and sometimes occur between these two parts of the verb unit (*phone the girl up*). They may also learn that a short list *him*, *her*, *it*, *us*, etc. always go between the verb and the particle.³

³ More probably, however, they learn a stress correlate for the construction. Bever et al. make a mistake about English struc-

Sentences 20–23 are not apposite for several reasons. First, Bever et al. seriously misunderstand my proposals: They assume I postulate that if a list of items occurs in a certain position, and one member of the list occurs in some nonhomologous position, then the entire list will occur in both positions. This is a stronger generalization mechanism than I proposed. Second, the anomaly in 23 presumably lies in the verb phrase *was cost*, not in the occurrence of *ten dollars* as subject. Third, my article specifically excluded the learning of transforms, so that neither 22 nor 23 are appropriate examples; my proposals could capture the structure of passives only insofar as they are similar to predicate phrases with adjective heads (cf. *The kazoo was blown by the child*; *The coffee was hot on the stove*).

Since it appears that Bever et al. have misunderstood the proposals, I must try to restate them more clearly and clarify their limitations. The discussion includes some recent thinking. The proposals are that what is learned are the temporal positions of units in verbal arrays and contingencies between morphemes. The position learned is the position of a unit within the next-larger containing unit of a hierarchy of units. Position within a unit may be defined absolutely (e.g., first) or relatively to a reference point (e.g., before *f*, first after *f*, second after *f*, where *f* is some frequently occurring morpheme—learning of such positions is shown in Braine, 1965b). Contingent morphemes need not be adjacent. All these positional and contingency relationships assumed to be learned could be represented by rewriting rules; for

ture when they assert that "it is a necessary condition upon the grammaticality of sentences of the form:

28. John phoned up X.

that X not be a pronoun [p. 474]." Pronouns can occur after the particle when they are strongly stressed (*THAT woman! John phoned up HER!*). There appears to be some requirement that the terminal unit in this construction have at least secondary stress; pronouns are not normally as strongly stressed, and this may be the reason the terminal placement is avoided.

example, the rules $X \rightarrow fAB$, $X \rightarrow gAC$, $A \rightarrow a_1, a_2, \dots$, $B \rightarrow b_1, b_2, \dots$, $C \rightarrow c_1, c_2, \dots$ represent that X phrases are prefixed by either f or g , that the a_i occur in second position, and that the b_i and c_i occur next-but-one after f and g , respectively. An adequate notation would presumably be a phrase-structure system of some kind. (The insufficiency of a finite-state notation is shown in Braine, 1965a.) A phrase-structure grammar is essentially a system for representing hierarchically organized positional relationships, and my proposals posit that it is these positional relations that are learned.

Although the theory appears to provide a sufficient basis for the learning of the treelike structures of a phrase-structure grammar, its handling of the learning of word classes requires examination. The theory implies some special limitations on the learning of unmarked word lists (i.e., word lists which are not correlated with particular function morphemes). The learner should have two kinds of difficulties with unmarked word lists: registering that (a) two such lists occurring at homologous locations are different lists, and that (b) one list occurring at two nonhomologous locations is the same list. The nature of the problem may be illustrated by considering a pair of rules of the form $X \rightarrow AB$, $X \rightarrow CD$, where A , B , C , and D are arbitrary lists of words. Exposed to such a system, the learner should readily learn which words go first and which second, but he has no way of distinguishing A items from C items, or B items from D items; thus, he should confuse the subsystems. (There is laboratory evidence for such confusion—Smith, 1965.) For the subsystems to be distinguished, either the classes have to be marked (e.g., $X \rightarrow A'B'$, $X \rightarrow C'D'$, $A' \rightarrow fA$, $B' \rightarrow gB$, $C' \rightarrow hC$, $D' \rightarrow kD$), or else the subsystems have to be marked (e.g., $X \rightarrow fAB$, $X \rightarrow gCD$, where the list positions are differentiated as "first after f ," "second after f ," etc.).

It follows from the above considerations that the kind of phrase-structure system for which my proposals are adequate is one which permits only one unmarked word list at any given location and which

restricts the circumstances in which an unmarked list can occur at two nonhomologous locations.⁴ A property of this sort may be called an "indexing" restriction, since such a grammar provides, in effect, an overt indexing system for its vocabulary lists. An indexing restriction would explain the purpose of function morphemes: Perhaps the human subject needs function morphemes to index classes and distinguish their locations. It is important to observe that a phrase-structure model which introduces classes as arbitrary lists of words (as in rules like $N \rightarrow \text{man, ball, } \dots$, $V \rightarrow \text{hit, take, } \dots$), and which imposes no restrictions on how the lists may be distributed in the rules, assigns to the human subject an *unlimited* capacity to learn and retain long lists of unrelated items occurring in complex positional relationships to each other without becoming confused. At the most general level, the question involved here is whether characteristics of human learning and long-term memory place any essential limitations on the form of a grammar. No limitations are built into the context-free and context-sensitive models (Chomsky, 1963) assumed by Bever et al., and these models do not distinguish function morphemes and lexicon. I would urge that laboratory studies of the learning of semantically empty systems can make an important contribution to the investigation of subjects' abilities to handle lists of unrelated words. A promising method for such studies is suggested elsewhere (Braine, 1965a; cf. also Smith, 1963, 1965).

Covert Categories

Despite the plausibility of some kind of indexing restriction, it is a fact that languages contain structures in which the same position can be occupied by two or more classes which are not differentially marked. Bever et al. are correct

⁴ To state this restriction precisely, much more experimental work is required: Locations can be partially homologous, and the circumstances in which lists occurring at partially homologous locations tend to be learned as the same list, or as different lists, have to be determined by experiment.

in pointing out that this sort of structure poses a problem for my proposals. These "covert categories," as they were called by Whorf (1956), may ultimately be based on semantic properties or relations. This notion would be consistent with the thinking behind my theory: Class markers provide a device for making heterogeneous items similar (they all "go with" the marker); however, when items are already similar by virtue of a semantic property or relation of some sort, it would be readily understandable that the overt marker might not be needed—the covert semantic marker would suffice to index the list. Covert semantic markers undoubtedly exist. Proper names provide an obvious example: A speaker hearing an exotic name for the first time presumably does not need to hear it used in noun positions in sentences in order to know its grammatical category. Unfortunately, semantic correlates are rarely so transparent. More often, although the class has a set of nuclear members which clearly have a common property, it also contains many words without the property. For example, Whorf (1956) says of some covert noun subclasses in Navaho:

Some terms belong to the round (or roundish) class, others to the long-object class, others fall into classes not dependent on shape. . . . I doubt that such distinctions are simply . . . recognitions of . . . objective differences. . . . One must learn as a part of learning Navaho that "sorrow" belongs in the round class [p. 91].

Since the speaker has to learn not only that round versus long is the basis of classification but also that "sorrow" is to "count as" round, it follows that a semantic-correlate theory of covert categories gives the correlate the status of a mediator in the sense of Jenkins and Palermo (1964).

Bever et al. are mistaken in asserting that the ideas of Jenkins and Palermo are "simply a weaker version of Braine's [p. 475]." ⁵ Bever et al. point out that Jen-

⁵ Incidentally, it is inappropriate of Bever et al. to attribute to Jenkins and Palermo a belief that the employment of mediational paradigms affords "a breakthrough" in the

kins and Palermo have little to say about phrase structure, and the difficulties faced by mediation theory in explaining position learning have been indicated by Smith (1965). However, Jenkins and Palermo do provide a reasonable theory of word-class formation that is at least theoretically capable of explaining covert categories. Their ideas are complementary to mine and not a weaker competitor. It is important to note, however, that my proposals imply a restriction on the scope of the four-stage stimulus equivalence paradigm: It should not be demonstrable without some semantic support for the equivalence classes.

SECONDARY ISSUES

Transformations and "Sublanguages"

My article was mainly concerned with phrase structure, and the brief overview on transformations was not intended as a precise statement about English, but rather as a general line of thought as to how a quest for what is learned might approach these aspects of English structure. English clearly does present a number of parallel phrase and sentence forms of the kind described as sublanguages. What might be learned in learning transformations is difficult to discuss in part because the transformation concept itself seems unclear. The uncertainty appears in at least three ways. First, it seems impossible to deduce from Chomsky's (1961) definition of a transformational grammar what features should be built into a verbal learning experiment in order to ensure that a subject is faced with the task of learning a transformation; laboratory study is therefore difficult. Second, it seems likely that a resolution of Putnam's (1961, p. 41) criticism, that Chomsky's definition of a transformational grammar is "much too wide," will affect the character of such grammars. Third, there are serious unresolved difficulties associated with the phrase structure of the products of transformation (Bach, 1964, pp. 70-78; Chomsky, 1955); discussing these, Chomsky (1962)

study of syntax. This term was not Jenkins and Palermo's.

has said candidly: "All answers I have seen so far are ad hoc. . . . I think . . . there is some important insight that is still lacking [p. 158]." Transformations are surely destined to be further clarified, and detailed discussion of their learning may have to wait upon the clarifications. Meanwhile, the difficulty of the laboratory study of transformations provides a pragmatic reason for concentrating on phrase structure.

Inflection

There is no truth whatever, nor warrant in my article, for the statement that I assume "that the processes employed in learning order languages are different in kind from those employed in learning inflected languages [p. 477]." Of course a theory of grammar acquisition should explain a child's ability to learn any language. However, I doubt that the Latin speaker learned what Bever et al. say he learned.

The major cues in Latin are based on relative position. The subject of the sentence is the element which occurs immediately before the nominative-case suffix; the object is the element immediately before the accusative-case suffix; the verb is the element to which suffixes comprising tense, mood, voice, aspect, and person morphemes are attached, etc. It is the learning of these kinds of relative positional relationships which has to be explained in giving an account of the learning of the grammar of an inflected language like Latin. It is not apparent why my viewpoint would have more difficulty with Latin-like languages than with English.

The analysis of Latin provided by Bever et al. leads to curious problems. If the "underlying strings" generated in the phrase structure have a rigid word order, then it must be assumed that the learner learns this order, even though no such order is visible in the language. It is difficult to see how the learner could learn this, or why one should want to postulate that he does. The analysis is not made more plausible by assuming that the learner also learns permutation rules governing the movements of the contents

of various positions to other positions. Moreover, the suffixes, which are clearly the cues defining the major grammatical positions, enter only peripherally into the formulation. I know of no generative grammar of Latin that has yet been published, and it may not even be certain that a feasibly simple grammar can be written in the manner suggested by Bever et al. Simple Latin sentences can contain more than three words. Six words have 720 possible orders, and eight have 40,320; even allowing for the rarity of some of these possible orders, a very large number of permutation transforms would presumably be required merely to generate the kernel sentences.

Intonation

I did not claim "that the discrimination of acoustic properties of speech signals accounts for the learning of segmentation [p. 479]." I suggested only that intonational cues might be helpful, a possibility Bever et al. admit.

Experimentation

With regard to my Experiment III, it is true, as several people have pointed out to me, that the segmentations are not inherent in the "languages" but are determined by the experimental procedure. Throughout the training phases, A and PQ phrases (in the A+PQ system) are presented as stimulus units in sentence-completion problems of the form $a_i \dots (a_j, p_i q_i)$ and $\dots p_i q_i (a_i, p_j q_j)$; it is assumed that this presentation of the phrases as units during training is sufficient to define the sentences as having two parts, an A phrase and a PQ phrase. The results of Experiment II provide some support for this assumption.

The reconstructive-memory technique (Braine, 1965a) seems superior to the sentence-completion procedure as a method for studying grammar acquisition. An important reason for dissatisfaction with the sentence-completion procedure is not that it fails to specify the segmentation, but rather that it *does* specify it. It is greatly preferable that the subject be required to discover the segmentation for himself. I have little doubt that the learn-

ing of positions within phrases can be demonstrated to the satisfaction of Bever et al.

Bever et al. raise objections to laboratory learning studies. However, their thought that young children may have especially rich capacities for grammar acquisition ("pretuned") is relevant only if it is assumed that these capacities do not operate in the laboratory; people who have speculated about the possibility of such capacities seem to believe that they endure until adolescence (Bellugi & Brown, 1964, p. 139), well past the age when laboratory work becomes feasible. I am unable to follow the second objection, that useful laboratory work is extremely difficult with artificial languages consisting of a finite set of strings. Their point that "any such language can be enumerated by a simple list [p. 481]" would be appropriate only if subjects learned a list of sentences as a list. On the contrary, memory for a list of patterned strings is primarily a reconstructive process and not a learning of a large number of individual items (Braine, 1965a; Smith, 1965). Subjects learn regularities in pattern that the strings exhibit, and "recall" is in large part a reconstruction of strings which have the regularities that the subject has registered. Thus, laboratory studies yield information about the pattern properties or "structural descriptions" learned.

In conclusion, I would like to return to what seems to me the fundamental issue at stake in this interchange: whether the order of elements in simple sentences reflects the order in the underlying string. It seems a very natural psychological approach to grammar to regard grammar acquisition as a learning of the patterned regularities in sentences, and my proposals are initial suggestions as to the bases of some kinds of regularities. This approach promises to be fruitful. However, Bever et al. are arguing that this approach is precluded because the learner is often never exposed to the pattern properties which must be assumed to be learned. Before accepting this argument, I would urge that psychologists examine the linguistic evidence for a distinction

between manifest and underlying structure in kernel sentences very closely indeed.⁶

⁶ (This footnote was added in proof following receipt of Bever et al.'s (1965a) rejoinder.) I see that Bever et al. now allow that it would be possible to write a grammar of English in which there would be no distinction between manifest and underlying structure in simple declaratives but claim that such a grammar would involve an unnecessarily complex system. They seem to assume that I would agree, since their section, "The Moral," accuses me of "playing fast and loose with simplicity constraints [p. 497]." I think that it is simply untrue that a phrase-structure analysis permitting discontinuous rules must lead to the kinds of complexities they claim: They do not appear to have investigated the potentialities of such an analysis thoroughly enough. Specifically, while they have seen that the affix-permutation rule (Chomsky, 1957, p. 39, Rule 29-ii; 1962, p. 141, Rule 12) would no longer be needed in the generation of simple declaratives, they have failed to see that it would not be needed elsewhere in the grammar either. This rule is the one which introduces the distinction between manifest and underlying structure into the auxiliary verb phrase, and the possibility of its total elimination (a gain in "simplicity" in the transformational part of the grammar) is one of the reasons for thinking that English contains discontinuous rules. Bever et al. say: "Of course the facts might have shown that the discontinuous phrase-structure rules are the appropriate linguistic analysis. To show this, the facts would have to be (at least) *that no sentence constructions require a transformational rule for affix movement* which can also be utilized by the simple declarative [p. 496; italics mine]." Since the italicized clause is true, the facts argue that discontinuous rules are *an* appropriate analysis. (Incidentally, there are independent objections to the affix-movement rule—see Bach, 1964, p. 81. The rule is not well formed because the symbol, "Affix," is introduced ad hoc, and reformulation is not simple.)

A word on "simplicity": It is possible to agree with Bever et al. that simplicity is an important consideration in scientific theorizing and still to reject their use of simplicity arguments. I am troubled by two main difficulties of principle.

1. Detailed measurement of simplicity (e.g., by means of a rule count, or other metric)

REFERENCES

- BACH, E. *An introduction to transformational grammars*. New York: Holt, 1964.
- BELLUGI, URSULA, & BROWN, R. (Eds.) *The acquisition of language. Monographs of the Society for Research in Child Development*, 1964, 29 (1, Whole No. 92), 1-192.
- BEVER, T. G., FODOR, J. A., & WEKSEL, W. Is linguistics empirical? *Psychological Review*, 1965, 72, 493-500. (a)
- BEVER, T. G., FODOR, J. A., & WEKSEL, W. On the acquisition of syntax: A critique of "contextual generalization." *Psychological Review*, 1965, 72, 467-482. (b)
- BLOCH, B. Studies in colloquial Japanese: II. Syntax. *Language*, 1946, 22, 200-248.
- BRAINE, M. D. S. On learning the grammatical order of words. *Psychological Review*, 1963, 70, 323-348.
- BRAINE, M. D. S. The insufficiency of a finite state model for verbal reconstructive memory. *Psychonomic Science*, 1965, 2, 291-292. (a)
- BRAINE, M. D. S. Learning the positions of words relative to a marker element. Washington, D. C.: 1965. (Mimeo) (b)
- BROWN, R., & BELLUGI, URSULA. Three processes in the child's acquisition of syntax. *Harvard Educational Review*, 1964, 34, 133-151.
- CHOMSKY, N. Transformational analysis. Unpublished doctoral dissertation, University of Pennsylvania, 1955.
- CHOMSKY, N. *Syntactic structures*. The Hague: Mouton, 1957.
- CHOMSKY, N. On the notion "rule of grammar." In R. Jakobson (Ed.), *Structure of language and its mathematical aspects, Proceedings of the 12th Symposium in Ap-*

must presuppose a particular form of grammatical model—it is doubtfully meaningful to count rules which are not of the same type. Competing grammars which differ in form (i.e., in the kinds of rules permitted) cannot therefore be compared in simplicity, except at the gross "intuitive" level that is usual in comparing scientific theories. While some very general questions (e.g., whether a phrase-structure grammar is adequate for the entire language) can be settled by simplicity considerations, because the differences in simplicity in the resulting grammars are great, it is hard to see that more detailed questions about the kinds of rules used in natural languages can be settled in this way. It follows that Bever et al. never confronted my original methodological point; if the component parts of a grammar are not separately responsive to data, and if simplicity considerations become inconclusive beyond a certain level of detail, then how are questions about the form of grammars settled, if not by convenience?

2. There is a major question about the domain over which simplicity should be taken. Should the domain comprise only the taken. Should the domain comprise only the possible grammars of a language (as Bever et al. assume), or should it also include possible accounts of the acquisition of the grammars? If there is a possibility that the simpler of two possible grammatical solutions might require the more complex acquisition theory, then the domain over which simplicity is taken cannot be restricted to grammar alone and must include acquisition theory—otherwise the grammarian merely

purchases simplicity at the psychologist's expense. The difficulty is well illustrated in Bever et al.'s main argument: They contest my proposals on the ground that simple English sentences contain a distinction between manifest and underlying structure; however, they acknowledge that a grammar could be written in which there is no such distinction, but claim that such a grammar should be rejected on grounds of complexity. Nevertheless, they allow that the grammatical solution containing the distinction might require a more complex acquisition theory (it would pose "serious problems for any theory of syntax learning"—1965b, p. 476). Thus, even on their own calculations, they cannot claim a *net* gain in simplicity for their analysis. In general, simplicity arguments are inherently question begging when they are used to constrain theories which are outside the domain over which simplicity is taken.

If the existence of a distinction between manifest and underlying kernel structure is to be settled by simplicity calculations (which seems undesirable), then surely the gain in simplicity accruing to an acquisition theory for taking the underlying structure as overt in simple sentences would likely be so great that it would more than compensate for an increase in grammatical complexity involved in writing the grammar so that the underlying structure was overt. And it has not been shown that there must be an increase in grammatical complexity. (Ultimately, it would seem parsimonious to identify grammatical simplicity with ease of acquisition.)

- plied Mathematics*. Providence, R. I.: American Mathematical Society, 1961. Pp. 6-24.
- CHOMSKY, N. A transformational approach to syntax. In A. A. Hill (Ed.), *Third Texas conference on problems of linguistic analysis in English*. Austin: Univer. Texas Press, 1962. Pp. 124-158.
- CHOMSKY, N. Formal properties of grammars. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology*. Vol. 2. New York: Wiley, 1963. Pp. 323-418.
- HARMAN, G. H. Generative grammars without transformation rules: A defense of phrase structure. *Language*, 1963, 39, 597-616.
- HARRIS, Z. S. *Methods in structural linguistics*. Chicago: Chicago Univer. Press, 1951.
- JENKINS, J., & PALERMO, D. Mediatonal processes and the acquisition of linguistic structure. In Ursula Bellugi & R. Brown (Eds.), *The acquisition of language. Monographs of the Society for Research in Child Development*, 1964, 29 (1, Whole No. 92), 141-169.
- MATTHEWS, G. H. Discontinuity and asymmetry in phrase structure grammars. *Information and Control*, 1963, 6, 137-146.
- POSTAL, P. Constituent structure: A study of contemporary models of syntactic description. *International Journal of American Linguistics*, 1964, No. 30. (Monogr. Suppl. Part III)
- PUTNAM, H. Some issues in the theory of grammar. In R. Jakobson (Ed.), *Structure of language and its mathematical aspects, Proceedings of the 12th Symposium in Applied Mathematics*. Providence, R. I.: American Mathematics Society, 1961. Pp. 25-42.
- SMITH, K. H. Recall of paired verbal units under various conditions of organization. Unpublished doctoral dissertation, University of Minnesota, 1963.
- SMITH, K. H. Mediation and position learning in the recall of structured letter pairs. *Psychonomic Science*, 1965, 2, 293-294.
- WELLS, R. Immediate constituents. *Language*, 1947, 23, 81-117.
- WHORF, B. L. *Language, thought, and reality: Selected writings*. Cambridge, Mass.: Technology Press, 1956.
- YNGVE, V. A model and an hypothesis for language structure. *Proceedings of the American Philosophical Society*, 1960, 104, 444-466.

(Received May 17, 1965)

IS LINGUISTICS EMPIRICAL?

T. G. BEVER, J. A. FODOR, AND W. WEKSEL

Massachusetts Institute of Technology

This paper continues the discussion of issues raised by Braine's theory of "contextual generalization." The arguments for analyzing the English declarative as transformationally generated are discussed at length. Broader issues about the nature of confirmation of claims made by grammars are also considered. It is argued that while the direct experimental verification of such claims is often not feasible, considerations of simplicity and generality can provide adequate grounds for their empirical confirmation or disconfirmation.

Braine's (1965) reply to our (1965) paper successfully clarifies the two major issues between us. The first of these concerns a detail of the grammar of English: Braine claims (and all generative grammarians have denied) that the simple declarative sentence has a special linguistic character in that it has no transformations in its derivational history. The second issue concerns the possible bases for confirming or disconfirming particular claims about the character of the grammatical analysis of a class of sentences. In our critique we argued that the special status Braine seeks for simple declaratives can be obtained only at the price of the ad hoc decision to treat them in isolation from the rest of the types of sentences in the language. This is equivalent to saying that a grammar which treats declaratives as nontransformational is inherently more complex than a grammar which does not. The methodological issue between us is that Braine does not accept such arguments as supplying even *prima facie* confirmation of the transformational characterization of declaratives. Braine apparently holds that each decision in the construction of a grammar must be susceptible to direct experimental confirmation or else be considered a mere artifact of the linguistic description, adopted "at the convenience of the linguist." It seems to us that Braine is certainly wrong on both points. On one hand the grammar that Braine proposes for simple declaratives is thoroughly unacceptable; on the other hand to accept Braine's view of confirmation would be to

make linguistic science (or, for that matter, any other kind of science) impossible. We shall discuss these points in reverse order.

Progress in linguistic analysis of natural language has depended on the careful separation of the theory of the language ("langue" or "competence") from the theory of the use of language ("parole" or "performance"). In this way the linguist has insulated himself from the fact that the variables determining the character of speech behavior reflect features other than the formal structure of the spoken language. For instance, the fact that some sentences are difficult to say, to remember, or to understand is obvious. Evidently, such facts are the consequence of interactions between linguistic variables and variables of memory, perception, motor integration, etc. To fail to so represent these facts would render impossible the representation of *either* the systematic character of language *or* the systematic character of speech behavior.

To avoid this consequence, linguistic analysis does not accept as pertinent to the characterization of *language* all facts that are pertinent to the characterization of verbal behavior. Only certain types of information about language are considered, for example, whether a sequence is intuitively grammatical, its relations to other sentences, the units of which it is constructed, etc. Such information is quite real and as palpable as any other psychological data; that *all* facts about speech are not considered in linguistic analysis does not reduce the scientific ob-

jectivity of those that *are* considered. A theory that can explain the latter phenomena is in fact a sound, *empirically supported* part of the total theory of language behavior. The linguist and psychologist seek to discover how the theory of the language, which explains one restricted set of data, is embedded within the theory of the speaker of a language, which explains speech behavior in general.

It is not at all surprising that the analysis of speech behavior should proceed from two empirical and theoretical sources. Indeed, distinguishing among the different kinds of data that constitute superficially homogeneous phenomena is absolutely universal in scientific explanations; it occurs wherever considerations of simplicity and explanatory power require that the observations be represented as interaction effects. Consider, for example, the analysis of a block sliding down an inclined plane. There are two kinds of variables that interact to determine the block's behavior—first, the forces acting downward on the body and determining the acceleration for an ideal system; second, the reactive forces (e.g., friction) due to the character of the particular body and plane under study. The observed behavior is susceptible of systematic explanation only on the view that it is the product of interactions between these distinct systems.

Since the theory of verbal performance is directly concerned with the behavior of speakers, it may often be subject to fairly direct experimental examination. The theory of competence, on the other hand, is concerned with the formulation of the linguistic information underlying verbal behavior. The speaker's competence is only reflected in his behavior via the kinds of performance variables mentioned above. The direct experimental verification of the theory of competence is correspondingly difficult.

This is not to say that the theory of competence is in any sense conventional or arbitrary. For, its support rests not only on occasional experimental confirmation but also on considerations of theoretical simplicity and power, fruitfulness, availability for integration with theories of performance variables, and so on. It

must be reemphasized that this relative inaccessibility to experimental manipulation is not specific to theories of linguistic competence. On the contrary, it is obvious that the more a scientific theory concerns itself with the fundamental mechanisms underlying the observables, the less susceptible it is to direct experimental test.

SOME EMPIRICALLY BASED FORMAL DECISIONS

We now consider some of the points where Braine feels that the form of the grammar is at the grammarian's (and his) arbitrary disposal. In each case we try to show that the particular claim about grammar we made in our original critique is responsive to some compelling facts about natural language. In any given instance it is, of course, possible to argue that the facts we invoke are wrong or that they might be accounted for in some more economical fashion. But it is *not* possible to claim that the form of the grammar is arbitrary. We will take the points in the order they appear in Braine's reply and omit discussion of Braine's new data or proposals and of the accuracy of our reading of his original paper.¹

THE NONPRIVILEGED STATUS OF SIMPLE DECLARATIVE SENTENCES

We argued that declarative sentences have neither linguistic nor psychological preeminence and that the underlying constituent structure of simple declarative sentences is itself an abstract object, just as in the case of the passive and all other constructions. Braine counters with several arguments indicating that we did not suggest evidence for this and that the only evidence he can think of relies on

¹ If we misunderstand Braine (1963) on some points in our critique, we apologize to him and to the reader and suggest that they both satisfy themselves that our misunderstandings were based on misreading. In this paper we quote Braine directly whenever possible. We do not utilize any linguistic discoveries made after 1957; recent theoretical modifications do not affect Braine's theory nor its incorrectness.

arbitrary decisions about the form of grammar:

[1.] . . . their long discussion of the passive transformation . . . is relevant only to the question of what the *passive* transformation is a transformation of. . . . This argument says nothing about the relation between the kernel grammar and simple declarative sentences [p. 484; emphasis ours].

2. The only argument he can recall in favor of an abstract underlying constituent structure (UCS) for simple declaratives is the treatment of discontinuous elements; he claims this is an artifact of an arbitrary decision about the form of grammar.

"Whether or not a phrase-structure (kernel) grammar can permit discontinuities is a *technical* question about the form of such grammars [p. 484; emphasis ours]." In particular, our claim that such constructions as *be chasing* are evidence for the distinction between manifest and underlying order is spurious: "These are of the $\lambda \dots \mu$ form [p. 485]." Braine suggests earlier that such discontinuities can be treated by a phrase-structure rule, $X \rightarrow \lambda \dots \mu$, where it is a formal convention that μ permutes with the next rightmost element. For example, this sequence of phrase-structure rules would give the phrase *be chasing* as an output of the underlying phrase structure:

$VP \rightarrow \text{aux } V$; $\text{aux} \rightarrow \text{be} \dots \text{ing}$; $V \rightarrow \text{chase}$.

A grammar which forbids representing the constituents in the form *be} \dots \text{ing} \dots* must necessarily generate [the order] *be} + \text{ing} + \text{chase}* [and we presume John + sg + pres + run] and . . . having generated the items in the wrong order, it is of course necessary to permute them into the actual order. *The entire argument associated with [these structures] rests on a particular choice of phrase-structure model* [p. 485; emphasis ours].

3. Braine then laments the alleged lack of basis for any decisions between "competing" phrase structure and transformational solutions of descriptive problems in syntax.

If the terminal strings generated by the phrase structure are permitted to be arbitrarily different from any actual sentence

structures, *there are no independent data against which the phrase structure and the transformational rules can be separately tested* . . . the grammarian can write the phrase-sentence kernel *partly on the basis of his convenience*, free to correct any poor fit with the manifest structure of the language by using transforms to reshuffle elements. This methodological looseness makes it impossible to accept empirical claims about the properties of phrase-structure grammars of natural languages [p. 485; emphases ours].

4. Braine claims, in summation, that we did not show that the simple declarative sentence has a distinct abstract form underlying it, and he asserts that to show it we . . .

. . . would have to show that there are no feasible forms of phrase-structure model which would permit transformational grammars to be written so that only the identity transform intervened between the phrase structure and the morphophonemic rules [p. 485].

Our replies can in some instances be quite brief.

1. Braine's argument on page 484 is simply wrong. On grounds of simplicity, *any* argument which shows that the underlying constituent structure for one sentence type is generated by a particular kind of phrase-structure rules is *prima facie* an argument that all sentence types have such rules underlying them. A grammar does not treat certain sentence types to the exclusion of others, and there is no reason to believe that a child does either.

2. It is difficult to understand what Braine can mean by the claim that the explanation of discontinuities is a "technical question." He correctly assesses the literature as showing that linguistic formalisms could provide either phrase-structure or transformational solutions. When nonequivalent explanations of the data compete, we must surely decide between them by asking which formalism accounts for the facts under consideration with the greatest economy and generality. If this makes the question of how discontinuous constituents are to be explained "technical," then *all* questions of scientific theory are technical, since considerations

of simplicity and generality underlie the solution of all such questions.

It is true that the use of discontinuous phrase-structure expansions could adequately generate number agreement and appropriate placement of the affix *ing* in the simple declarative sentence. But that is *all* it could do. The arguments in our critique showed that for other constructions (e.g., passive, negative passive, question passive, cleft-sentence passive, negative-question passive, etc.) the affix-movement rule must *follow* the passive rule, which itself follows the development of the base structure. Thus, the affix-movement rule needed for these constructions *cannot* be in the phrase structure and is consequently a transformation. Indeed, since the same transformation is formally capable of accounting for number agreement and affix permutation in the simple declarative as well as other constructions, we have a clear choice between two types of grammar. One grammar uses a discontinuous phrase-structure rule for number attachment in the declarative and a transformation for number attachment in all other constructions. The other grammar employs only the transformation. Clearly the latter grammar must be preferred since it explains the facts with less duplication and with a more restricted form of grammar.

Of course the facts *might* have shown that the discontinuous phrase-structure rules *are* the appropriate linguistic analysis. To show this, the facts would have to be (at least) that *no* sentence constructions require a transformational rule for affix movement which can also be utilized by the simple declarative. That is, the transformational solution *could* have been invalidated by the character of English, and thus the decision to reject the discontinuous phrase-structure solution is not based on the "grammarian's convenience," but on the facts of the language.

3. It is hard to imagine what the "independent data" might be that could "separately" confirm any linguistic analysis since the theoretical decisions themselves are based on all presently available empirical considerations extractable from the language.

Perhaps we can additionally support the decision to treat number agreement (and affix-movement phenomena) with one transformation by an appeal to the reader's intuition about some previously unconsidered facts. Is it not the case that the subject-verb number agreement (and formation of *be + ing*) is the same sort of relation in these two sentences?

They were running.

They were being served.

The analysis proposed by Braine generates these two subject-verb agreements by two distinct *kinds* of processes,² and we might expect that if this analysis were true, different intuitions about the nature of the agreements could be informally or experimentally extracted from native speakers. In fact, we have no reason to believe that such differences exist. Number agreement seems to be a psychologically unitary process in English. This is reflected in the formal analysis to which the linguistic facts led us.

4. In short, although there *is* probably a "feasible" form of grammar in which the underlying and derived constituent structures of simple declaratives would be identical, this analysis is blocked for English on empirical grounds. We have shown briefly that at very least such a solution would involve an unnecessarily complex system since there is a simpler analysis which accounts for the relevant facts. In this way we show that Braine's claim that there is a "base" or "kernel" grammar producing all and only simple declaratives to be *logically* possible but empirically unacceptable.

THE MORAL

We have argued that there is strong empirical support for the claim that declaratives have abstract underlying structures and that this support derives primarily from considerations of the simplicity and generality of grammatical rules. In conclusion it may be pointed out that, though it is true that a suitable complication of the grammar would permit Braine to treat simple declaratives

² Discontinuous phrase structure for the first, transformational for the second.

as the only untransformed structures, it must be added that similar complications would permit that treatment for any other sentence type. There is no more reason for complicating the grammar in order to render declarative sentences uniquely nontransformational than there is for complicating the grammar in order to render passives, questions, negatives, imperatives, etc. uniquely nontransformational. The difficulty with playing fast and loose with simplicity constraints is that, once having started, it is hard to find a way to stop.

THE EXCLUSION OF TRANSFORMATIONS

Braine states that the "... main purpose [of my discussion of transformations] was to show that [their] existence did not contradict my proposals concerning simple declaratives [p. 484]." And later (p. 486), "my [proposal] specifically excluded the learning of transforms . . . my proposals could capture the structure of passives only insofar as they are similar to predicate phrases with adjective heads (cf. *The kazoo was blown by the child; The coffee was hot on the stove.*)" Furthermore he states that the fact that many transforms appear similar to simple declarative sentences explains why it is reasonable to assume that simple declaratives are learned separately by the child, even though there is no evidence that they preponderate in the child's verbal environment:

I did not claim that children were mainly exposed to simple declaratives. . . . Information about positional and contingency relations in simple sentences is abundantly exemplified in many transforms . . . consider the similarity in verb-phrase structure between passives and sentences with adjective predators (e.g., *George was served—George was sensible*). . . . Experience with the verb structure of one sentence type could hardly fail to transfer to the other type [p. 483; emphases ours].

This line of argument appears to involve Braine in a hopeless dilemma. To wit, if the child cannot exclude transformations, then our objections in the critique and in this reply obtain; if the child can isolate and reject transforms

from consideration, then he must already know the grammar including its transformational component: Transformational information is required to distinguish between declaratives and other types of sentences.

This dilemma runs deep. If the child cannot yet tell whether an apparent declarative sentence actually is one, how does he know when to transfer experience from one putative declarative to another? It is undeniable that many sentences other than declaratives appear in the declarative-like form "NP VP." How does the child know which such sentences are sufficiently "like" true declarative to permit relevant transfer of training unless he already has transformational information? How can "experience with one type . . . transfer to another" sentence of the same type unless the child can see that they are of the same type? That is, he must have the transformational information required to distinguish between real and apparent declaratives. Furthermore, how does the child know what kind of experience to generalize? Surely if everything the child knows about

George was served.

generalized to

George was sensible.

the child should expect the sentence

**Somebody sensibled George.*

to parallel

Somebody served George.

THE DISCRIMINATION OF HOMOLOGOUS SEQUENCES AND "COVERT CATEGORIES"

Several different problems for Braine's earlier and present proposals are manifestations of Braine's inability to answer the question: How are sentence types differentiated when they have similar word-class orders? (See our critique for examples.)

Braine says, "... the kind of phrase-structure system for which my proposals are adequate is one which permits only one unmarked word list at any given location and which restricts the circumstances in which an unmarked list can occur at two nonhomologous locations [p. 487]." He proposes that this requirement be met

by an "indexing restriction" partially carried by "function words" which "index classes and distinguish their locations [p. 487]." ³ But, he notes, this is not an adequate solution: "... it is a fact that languages contain structures in which the same position can be occupied by two or more classes which are not differentially marked. Bever et al. are correct in pointing out that this sort of structure poses a problem for my proposals [pp. 487-488]."

As a solution, Braine invokes the use of "covert categories" which he presumes to be "semantic." Words that are otherwise undifferentiated in a particular sentence type thus must have an "internal" analysis which differentiates them.

We, of course, agree with this proposal since it reduces to the tautology that whatever *really* is different ought to be so described. But the tautology is useless unless a theory for providing correct internal analyses is also forthcoming, and Braine's own discussion suggests how unlikely it is that such analyses can be formulated in semantic terms. In fact, generative grammars do provide differing types of words in homologous position with distinct analyses; they do so largely by the use of transformations which can distinguish between items that have similar privileges of occurrence but which differ in the base structures from which they derive. Consider first a nontransformational case—proper names (an example proposed by Braine): Presumably there are different internal analyses for the lexical items "George" and "butter" which allow the sentences

George is nice.

Butter is nice.

and the sentence

The butter is nice.

but block the otherwise contextually generalizable

The George is nice.

³ Since Braine's first paper, rules for the introduction of lexical items into underlying structures have been shown to be substitution transformations (Chomsky, 1965; Matthews, 1963) and not phrase-structure rules. His argument at the end of his section, "The Scope of the Proposals," is therefore unsound.

But what internal markers differentiate the participle in these sentences?

Making mistakes can be annoying.

Recurring mistakes can be annoying.

What "markers" allow

Many mistakes were made.

but block

**Many mistakes were recurred.*

The analysis which differentiates the participles successfully must refer to the underlying constituent structure, namely to the fact that "make" occurs with an object permissible in the UCS (and that "mistakes" is that object) and that "recur" does not (so that "mistakes" is the subject in the UCS).⁴

In short, we agree with Braine's proposal that words and homologous sentence types are differentiated by characteristics *not directly observable in the sequences themselves*. In addition we propose that differences in transformational derivation are necessary to provide the explanation for many of these "covert" distinctions. So far as we can see, Braine has provided no reason whatever for rejecting this suggestion.

THE "UNCLARITY OF TRANSFORMS"

Braine states: "What might be learned in learning transformations is difficult to discuss in part because the transformation concept itself seems unclear [p. 488]." The "unclear" appears in at least two general ways: (a) "... it seems impossible to deduce from Chomsky's (1961) definition ... what features should be built into a verbal learning experiment in

⁴ Notice that this exactly characterizes the distinction between "transitive" and "intransitive" verbs, although this distinction is not apparent in the surface structure of the sample sentences. We could decide that such notions as *transitive*, *intransitive*, and so on should be represented as "covert features" of lexical items without reference to the UCS. This possibility is empirically rejected because it would not eradicate an abstract level from the grammar, and it would result in duplication in the surface phrase marker of the distinctions already made in the UCS. It is never difficult to construct a theory *more* complicated than the simplest one currently available.

order to ensure that a subject ... [learns] a transformation [p. 488]"; (b) grammars are too strong and their character will change as they become refined. Rules for derived constituent structure are not well understood.

Regarding (a), it seems to us fantastic to require that experiments should be *deducible* from theories. Regarding (b), the questions in grammatical theory which await theoretical and empirical elucidation do not place in jeopardy the answers which we do have.

ON LABORATORY EXPERIMENTATION WITH PRIMITIVE ARTIFICIAL LANGUAGES

In our critique we suggested that experiments on artificial languages are equivocal with respect to natural language. It is inconclusive to show that the subjects learn the intended structure when exposed to a relatively small number of instances either in artificial or natural language since the structure that must be assumed for natural languages is determined by the need to supply a simple, uniform treatment for a large variety of sentences of very specific formal character. Finally there is no reason to believe that simple phrase-structure artificial languages are on a psychological continuum with natural language.

Braine counters in his reply that he has other evidence indicating that the subjects did learn the assumed structure. He also points out that it is silly to assume that a child loses his natural language-learning capacity when he enters a laboratory.

Of course we were only questioning the *experiment's* capacity to bring out the child's ability. There is no support for the assumption that the experiment with this restricted "language" elicits the mechanisms employed in learning natural language. Evidence that these languages do not elicit first-language learning mechanisms can be found in the recent work of Wales and Grant.⁵ They repeated Braine's Experiment III with adults and obtained exactly the same results. This means either that adults learn languages in the same way as children or that the experi-

ment is not tapping the language ability indigenous to the child.⁶

THE CONFLICT BETWEEN LINGUISTIC ANALYSIS AND CURRENT LEARNING THEORY

In our critique, we suggested that current models of learning are in fact incompatible with the empirically based results of linguistic theory. Braine is sensitive to this, and he implicitly uses it as an argument *against* certain features of linguistic analysis; in particular, he is bothered by the concept of an "abstract" underlying constituent structure for which no explicit representation appears.

The distinction [between manifest and underlying structure in kernel sentences] is certainly not convenient psychologically: The assumption that the learner learns structures to which he is never exposed would indeed pose "serious problems for *any* theory of syntax learning [p. 486]."

... the fundamental issue at stake in this interchange [is] whether the order of elements in simple sentences reflects the order in the underlying string. ... Bever et al. are arguing that

... the learner is often never exposed to the pattern properties which must be assumed to be learned. Before accepting this argument, I would urge that psychologists examine the linguistic evidence for a distinction between manifest and underlying structure in kernel sentences very closely indeed [p. 490].

We of course agree, and we have tried to expand and clarify the empirical basis for the decision (a) that *all* sentences

⁶ It seems to us unlikely *on any account* that adults learn languages in the same manner as children do. This does not depend on an assumption of "innateness" for language learning in children, nor must it reject the possibility that "proactive inhibition" explains the adult's difficulty with language learning. Whatever view you hold, it is obvious that an adult with 20-years experience in Language X will learn Language Y differently from a child that has had only 2- or 3-years active experience in X. The fact that children and adults yield identical experimental results thus indicates that the experiment is not stimulating first-language learning mechanisms in children.

⁵ Wales, personal communication, 1965.

have an abstract structure underlying them and (b) that this structure is not isomorphic with the simple declarative sentence. The essential themes of our discussion have been: (a) Related parts of a language cannot be considered in isolation; (b) decisions concerning the form of grammar are based on determining which formalism explains the facts of the language with the least duplication and the most generality.

The psychological role of the underlying constituent structure has not been carefully studied except insofar as it is an integral part of the account of grammatical sentences themselves. Though linguistic evidence of this latter sort is not in any sense arbitrary or equivocal, we do feel that the exact form of the empirical extensions of the underlying constituent structure into other aspects of language behavior requires thorough psychological investigation. Part of that study will no doubt include experimental effects of UCS upon the recall, perception, or learning of sentences. A crucially important part of that study must center on *how* the child goes about learning structures for which there are no explicit models.

There are those who agree with Braine

that this is impossible either in fact or in principle. As the empirical basis for assuming an abstract underlying structure in language becomes broader and the explanatory power of that assumption becomes deeper, *we recommend to all psychologists that they seriously question the adequacy of any theory of learning that cannot account for the fact that such structures are acquired.*

REFERENCES

- BEVER, T. G., FODOR, J. A., & WEKSEL, W. On the acquisition of syntax: A critique of "contextual generalization." *Psychological Review*, 1965, 72, 467-482.
- BRAINE, M. D. S. On learning the grammatical order of words. *Psychological Review*, 1963, 70, 323-348.
- BRAINE, M. D. S. On the basis of phrase structure: A reply to Bever, Fodor, and Weksel. *Psychological Review*, 1965, 72, 483-492.
- CHOMSKY, N. Topics in the theory of generative grammar. In T. A. Sebeok (Ed.), *Current trends in linguistics*. Vol. 3. *Linguistic theory*. New York: Humanities Press, 1965.
- MATTHEWS, G. H. Discontinuity and asymmetry in phrase structure grammars. *Information and Control*, 1963, 6, 137-146.

(Early publication received July 19, 1965)

ECOLOGICAL OPTICS AND VISUAL SLANT¹

ROBERT B. FREEMAN, JR.

Pennsylvania State University

Flock's "A Possible Optical Basis for Monocular Slant Perception" is criticized as being a theory of stimuli rather than a theory of perception. To account for accurate monocular slant perception, the theory requires 9 assumptions, including the unproved ability of the eye to register random texture density. The alternative hypothesis is proposed that monocular visual slant is a function primarily of contour perspective which varies with the size, shape, and viewing distance, as well as slant, of plane surfaces.

Flock (1964a), in a recent paper, elaborates on Gibson's (1950, 1959, 1961) gradient concept which relates visual ecology to retinal stimulation. Assuming that "slant perceptions . . . depend on optic variables [p. 391]," Flock presents an analysis of "optic variables" arising from textured, slanted surfaces. The arguments of this note are that the principal part of Flock's theory is not a theory of perception but a description of visual stimuli, that the casually mentioned "abilities" of the eye to register the proposed visual variables of surface slants involve an inordinate number of assumptions for what they accomplish, and that the visual variables chosen for description in the paper have already been experimentally demonstrated to be ineffectual and unnecessary for the perception of slant. An alternative stimulus for visual slant is suggested.

ECOLOGICAL OPTICS OF SLANT

The purpose of Flock's paper is to show how

. . . *accurate* monocular slant perceptions are possible even though a motionless viewer has no previous experience with a particular substance, even though the textural elements of a motionless surface are irregular in size, shape, and separation, and even though parts of the surface are illuminated differently [Flock, 1964a, p. 380; italics, mine].

¹ This note was written in connection with research supported by Grant MH08856-01, National Institutes of Health, United States Public Health Service.

Flock attempts to accomplish this goal by describing a trigonometric transformation of surface elements which remains invariant with variations in stimulus conditions other than physical slant. His analysis is primarily an ecological description of the optics of slant: a description of the transformations which light reflected from real, planar surfaces undergoes in projection as it converges on the observer's eye.

His analysis is based on two postulates. Postulate I states that "substances" (presumably surfaces) of a certain class possess a unique pattern characteristic of that class. Postulate II states that such patterns consist of "like elements" which vary randomly in size, shape, and separation. Flock thus allows for variability in the distal stimulus, but the variability is counterbalanced by dividing the textured surface into linearly equal units of n like elements. Somehow it is argued that the larger the unit, and hence the larger the n , the greater the regularity.² It is not explained how the eye can register such a unit of n like elements from among a

² An equation is presented for "degree of regularity" of like elements which has some peculiar properties. Regularity is made an inverse function of the variability of element size but a direct function of both average size and total number of like elements. Equation 4 has the result that stimulus elements of very high variability can have a high degree of regularity if there are enough of them. Conversely, if the n is small compared to s , negative regularity can result.

continuously variable series of like elements.

Flock then attempts to make "optical slant" a function of the relative projective angles of three such linearly equal units which are separated by equal visual angles perpendicular to the axis of apparent rotation. This is where Flock's optical theory of slant encounters its greatest trouble. It can be easily shown that the angular subtense of any three distances of equal length parallel to the axis of rotation of a plane surface changes in a complex fashion with rotation of the surface. This fact suggests that the texture gradient at the eye arising from elements scattered about a plane surface will not be a simple one and will vary with size, distance, and other variables, as well as slant (cf. Freeman, in press—b).

But, assuming that slanted, planar surfaces are perceived as slanted, planar surfaces at their true slant, Flock argues that there must be something in the stimulus situation which yields such a veridical perception. His Equation 5 represents an effort to turn a complex stimulus into a simple one so that the eye can have a simple percept. Equation 5 expresses the variation in visual angle subtended by a single unit of "optical n " with variation of visual slant ("optical θ ") of the distal surface. Although the derivation of Equation 5 is not made available, it has the property that optical θ is invariant with the number of optical units involved in directions both parallel and perpendicular to the axis of rotation. Thus optical θ must be unrelated to the size of the plane surface. This requirement makes necessary a very complex equation of a form unknown to psychophysics. As will be shown below, such a complex function is unnecessary.

FLOCK'S PSYCHOPHYSICAL ASSUMPTIONS

Flock has argued for a potential visual stimulus for slant in the manner in which Gibson (1961) has presented a general description of "ecological optics" of the visual environment. But in presenting a potential visual stimulus, he must also postulate a potential sensory system to respond to the potential slant stimulus or

optical θ . To do so, nine assumptions are required (Flock, 1964a, pp. 382, 389), the last of which is that the visual system registers texture in a manner conforming to Equation 5. Among other things, the eye is required to choose an appropriate optical n for each optical unit according to an additional set of four rules (Flock, 1964a, p. 384). With 13 assumptions and criteria involved, Flock's *psychophysical* theory of optical slant is complicated indeed. No suggestions are made as to *how* the visual system might be expected to bring about such a complicated analysis.

NEGATIVE EXPERIMENTAL EVIDENCE

Aside from the theoretical complexities of optical θ there are several experimental objections to the theory. Flock (1964b) himself has already provided evidence that the optical registration of motionless texture is poor. In his Experiment IX, using the Gibson-type shadow caster and an electrostatic texture, subjects adjusted a protractor to indicate the apparent slant of a 69° field of texture gradients corresponding to nine different physical slants at 10° intervals from -40° (top toward the subject) to $+40^\circ$ (top away from the subject). Flock regrettably gives the results in terms of the regression of the subjects' judgments of slant on the slant of the shadow caster rather than plotting the former as a function of the latter in the usual fashion. But these data are sufficient to show the poor psychophysical correspondence involved, insofar as the mean regression coefficient in Experiment IX was only .13 as compared to 1.12 in Experiment VIII in which motion-parallactic cues were available with the identical stimulus situation. The gradient arising from random texture at a slant is an inadequate stimulus for visual slant for the motionless observer.

If texture gradients are not effective stimuli for slant, what other proximal stimuli are available? An earlier experiment (Clark, Smith, & Rabe, 1956) had compared the relative effectiveness of texture gradients and "outline gradient," or perspective. In their report, Condition B was a randomly textured surface slanted

at 40° viewed through a 6-centimeter reduction hole, while Condition C was a *textureless* rectangle (with the same solid-angle area as the hole in B) on a black background, also at 40° slant. Judged slant in C was significantly greater than in B. Judged slant in Condition E, which *combined* texture and outline perspective, was *not* significantly greater than in the textureless-rectangle condition. According to the results of Clark et al. (1956), outline perspective has a significantly greater effect on judged slant than texture gradients. And when combined, outline dominates as a visual stimulus. Another experiment (Gruber & Clark, 1956) varied both size and texture density of random-dot patterns as well as the distance of the observer. Observers greatly underestimated the three different surface slants used. Furthermore, both surface texture and observation distance had significant effects on judgments. These results are clearly inconsistent with the concept that "accurate monocular slant perceptions are possible even though . . . the textural elements of a motionless surface are irregular in size, shape, and separation . . . [Flock, 1964a, p. 380]."

Finally, prompted by a finding by Stavrianos (1945) that the judged slant of rectangles of constant shape varies with their size, this writer (Freeman, in press—a) conducted a parametric study to determine whether the size effect could be attributable to outline perspective. Textureless rectangles of sizes ranging from 8–40 centimeters in length were viewed monocularly under complete reduction conditions at a distance of 135 centimeters. With a 24-centimeter reference rectangle, equal-slant contours were obtained for five slants (15°, 30°, 45°, 60°, and 75°), both forwards and backwards. The size effect on judged slant was large and highly significant in most of the curves. Since outline perspective of slanted rectangles varies with the width and probably height, as well as physical slant, of the rectangles, the size effect appears to be a function simply of projective outline shape.

CONTOURS—PRINCIPAL STIMULI FOR SLANT

In addition to the behavioral studies mentioned above, there is mounting evidence that the vertebrate visual system is "tuned" to register contours (abrupt brightness gradients). In addition to the extensive investigations of Hartline and Ratliff (e.g., Ratliff, 1961) on the inhibitory interactions of the lateral eye of the arthropod *limulus*, there are the discoveries of neural boundary detectors in the optic tract of the frog (Maturana, Lettvin, Pitts, & McCulloch, 1960) and in the visual cortex of the cat (Hubel & Wiesel, 1962). The combination of behavioral and electrophysiological evidence cited above makes possible two general postulates which are in disagreement with Flock's optical texture gradient hypothesis:

1. With monocular observation, visual shape and visual slant are a function primarily of linear outline perspective.
2. The greater the numerical value of linear perspective, the greater the judged visual slant and the greater the effect on judged shape.

The second postulate implies that apparent slant will vary with perspective, whether or not perspective is a true function of physical slant, and with or without nonlinear (random) surface texture. Since outline perspective varies with the shape, size, and distance of a plane surface as well as with slant, the apparent slant of stimuli so varied, when viewed monocularly with a motionless head under complete reduction conditions, must also vary, Flock's optical *n* notwithstanding.

The contour hypothesis therefore does not predict veridical judgments of slant, in the sense of judgments which are consistent with physical rotation relative to a fixed observation position. On the contrary, it says that visual slant will vary with variation in the projective character of stimulus contours at a slant to the visual axis, regardless of the environmental source of such variation. It is an explanation based, not on optical ecology, but on retinal stimulation. It is, finally,

a theory of perception, not a theory of stimuli.

REFERENCES

- CLARK, W. C., SMITH, A. H., & RABE, AUSMA. The interaction of surface texture, outline gradient, and ground in the perception of slant. *Canadian Journal of Psychology*, 1956, 10, 1-8.
- FLOCK, H. R. A possible optical basis for monocular slant perception. *Psychological Review*, 1964, 71, 380-391. (a)
- FLOCK, H. R. Some conditions sufficient for accurate monocular perceptions of moving surface slants. *Journal of Experimental Psychology*, 1964, 67, 560-572. (b)
- FREEMAN, R. B., JR. Effect of size on visual slant. *Journal of Experimental Psychology*, 1966, 71, in press. (a)
- FREEMAN, R. B., JR. Function of cues in the perceptual learning of visual slant. *Psychological Monographs*, 1966, in press. (b)
- GIBSON, J. J. *The perception of the visual world*. Boston: Houghton Mifflin, 1950.
- GIBSON, J. J. Perception as a function of stimulation. In S. Koch (Ed.), *Psychology: A study of a science*. Vol. 1. New York: McGraw-Hill, 1959. Pp. 456-501.
- GIBSON, J. J. Ecological optics. *Vision Research*, 1961, 1, 253-262.
- GRUBER, H. E., & CLARK, W. C. Perception of slanted surfaces. *Perceptual and Motor Skills*, 1956, 6, 97-106.
- HUBEL, D. H., & WIESEL, T. N. Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex. *Journal of Physiology*, 1962, 160, 106-154.
- MATURANA, H. R., LETTVIN, J. Y., PITTS, W. H., & McCULLOCH, W. S. Physiology and anatomy of vision in the frog. *Journal of General Physiology*, 1960, 43, Suppl., 129-175.
- RATLIFF, F. Inhibitory interaction and the detection and enhancement of contours. In W. A. Rosenblith (Ed.), *Sensory communication*. New York: McGraw-Hill, 1961. Pp. 183-203.
- STAVRIANOS, BERTHA K. The relation of shape perception to explicit judgments of inclination. *Archives of Psychology*, 1945, 296, 94.

(Received October 29, 1964)

OPTICAL TEXTURE AND LINEAR PERSPECTIVE AS STIMULI FOR SLANT PERCEPTION¹

HOWARD R. FLOCK

Dartmouth College

Experiments critical of the effectiveness of variables of optical texture in evoking accurate judgments of slant are shown to be inappropriate, inadequate, or deficient. Experiments supporting linear perspective as a stimulus for slant are evaluated. The contention that perceived slant is a function primarily of linear perspective is shown to be oversimplified and hardly adequate to cope with the facts. Some aspects of Flock's theoretical model specifying optical stimuli for slant are discussed.

In his criticism of my theoretical paper on visual slant perception (Flock, 1964a), Freeman (1965) claims that variables of surface texture are both ineffectual and unnecessary for the perception of slant. He argues that all perceived slants are a function primarily of linear perspective (contour convergence) and that the greater the linear perspective the greater the judged visual slant. To support his argument he cites two classes of empirical evidence which seem to show that (a) the bounding contours of rectangles, not their in-lying texture, induce perceived slant (Clark, Smith, & Rabe, 1956a), and (b) surface texture by itself induces considerable underestimates of slant (Gruber & Clark, 1956; Flock, 1964b, Experiment IX), whereas textureless rectangles in isolation consistently induce rather good judgments of slant (Freeman, in press). In addition, Freeman apparently was dissatisfied with my theoretical model describing potential optical variables for perceived slant.

TEXTURELESS RECTANGLES VERSUS SURFACE TEXTURE

Clark et al. (1956a) found that surface texture by itself ($b = .23$; in 1956b, $.24$)²

¹ This work was supported by National Science Foundation Grant GB 2474.

² In order to make the results of different experiments comparable, I have expressed their results in the form of regression coefficients, b , taking the regression of mean judged slant over the " k " levels of physical slant employed by the experimenter (E).

is a less-effective condition for accurate slant judgments than is a rectangle by itself ($b = .34$ and $.30$; in 1955, $.42$ and $.44$). Moreover, adding texture to a textureless rectangle does not change the judgment of slant ($b = .36$). These studies are faithfully described by Dember (1961) who seems to accept the general conclusion in favor of outline convergence. More recently Smith (1964, Conditions A and B) replicated the earlier findings and again found that textured and textureless rectangles do equally well ($b = .67$ and $.68$).

Using Clark et al.'s (1956a) arrangement of apparatus, Gruber and Clark (1956) tested the effects on slant judgments of surface texture in isolation. They varied distance of surface, diameter of texture elements (white dots), and separation of texture elements. Regression coefficients based on extrapolations from a partial graphic report of their data varied between $.18$ and $.43$ depending on distance, dot density, and dot size.

More recently Epstein (1962), Epstein and Mountford (1963), and Freeman (in press) have found that small, textureless

In the many cases where E used just one physical slant, 0° slant was added gratuitously on the assumption that had the surface been at 0° slant, subjects' (S 's) judgments would have been centered on 0° . When there is an identity relation between the levels of judged and physical slant, $b = 1.0$, and the judgments are assumed to be optimal. When $b = 0$, judgments are at a chance level.

TABLE 1
ANGULAR PARAMETERS IN 10 EXPERIMENTS

Es and experimental conditions		Display surfaces	V (in deg.)	D (in m.)	E (in mm.)	S (in mm.)	θ (in deg.)	α (in deg.)	$\delta\gamma-\delta\epsilon$ (min. of arc)	$\delta\gamma$ (min. of arc)	$\delta\gamma-\delta\epsilon$ (min. of arc)	b
Clark et al.	1956a	B	10.4	1.63	6	11	40	—	—	13.57	11.63	.23
Clark et al.	1956b	B	10.4	1.63	6	11	40	—	—	13.57	11.63	.24
Clark et al.	1956a	C, D	20.2	1.63	—	—	40	9.8	36.	—	—	.34, .30
Clark et al.	1956a	E	20.2	1.63	6	11	40	9.8	36.	13.36	11.97	.38
Clark et al.	1955	T-R	20.2	1.63	—	—	0.20, 40	9.8	36.	—	—	.42
Clark et al.	1955	R	20.2	1.63	—	—	0.20, 40	6.7	24.	—	—	.44
Clark et al.	1955	R	20.2	1.63	—	—	0.15, 30, 45, 60	6.4	29.	—	—	.67
Smith	1964	A	32.5	2.25	6	13 ^b	0, 15, 30, 45, 60	6.3	29.	9.63	8.74	.68
Smith	1964	B	32.5	2.25	—	—	32, 43, 53	—	—	1.85	1.58	.24, .20
Gruber et al.	1956	I	7. ^a	6.	3	8 ^b , 16 ^b	32, 43, 53	—	—	3.71	3.15	.22, .23
					6	8 ^b , 16 ^b	32, 43, 53	—	—	3.71	3.15	.27, .25
					3	8 ^b , 16 ^b	32, 43, 53	—	—	2.47	2.10	.25, .19
					6	8 ^b , 16 ^b	32, 43, 53	—	—	4.95	4.20	.28, .29
					3	8 ^b , 16 ^b	32, 43, 53	—	—	7.42	6.31	.28, .34
					6	8 ^b , 16 ^b	32, 43, 53	—	—	7.42	6.31	.24, .21
					3	8 ^b , 16 ^b	32, 43, 53	—	—	14.84	12.61	.33, .43
					6	8 ^b , 16 ^b	32, 43, 53	—	—	16.73	12.88	.28
		II	15.9	1.38	6	6 ^b	43	—	—	16.73	12.88	.35
		T	—	—	—	30 ^b	43	—	—	16.73	12.88	.18
			—	—	—	90 ^b	43	—	—	16.73	12.88	.41, .44
Smith	1956	A, C	20.2	1.63	—	—	0, 10, 20, 30, 40, 50	9.8	43.	—	—	.35
Smith	1959	A	21.4	1.73	—	—	0, 10, 25, 40	9.	30.	—	—	.35
Epstein	1962	NR	?	1.5	—	—	15, 30, 45, 60	.97	.65	—	—	.59

^a Rectangular aperture.

^b Averages of variable parameters.

rectangles under complete reduction conditions yield almost perfect slant judgments. Epstein's experiments (Conditions NR) yielded coefficients of .59 and .69. Freeman's experiments yielded a total of 45 coefficients of regression of mean judged slant over 10 levels of physical slant, varying from .67 to 1.17 with a near-perfect mean and median of .99.

At the superficial level that I have described these data, it would seem that there might be some basis for Freeman's (1965) claim. But let us look more closely at these experiments cited by Freeman.

TABLE 1

Some relevant data about these experiments are given in Table 1. The experimenter (*E*) and experimental conditions are identified in Columns 1-3. The symbols T, R, and T-R in Column 4 indicate that the display was composed of an extended (edges occluded from view) textured surface (T), of a textureless rectangle (R), or of a textured rectangle (T-R). Columns 5-8 give for each experiment the total angular diameter of the field of view (*V*) in degrees, the distance (*D*) in meters between the center of the display and the eye, the diameter (*E*) in millimeters of the texture element, and the separation (*S*) in millimeters from center to center of adjacent texture elements. The symbol θ , in degrees, indicates the actual slants used in an experiment. The symbol α gives the angular height in degrees of the rectangle when at 0° slant, measured along the meridian perpendicular to the axis of rotation. The column labeled $\beta_N - \beta_F$ provides a rough measure of linear perspective, β_N being the angular width in minutes of arc of the slanted rectangular edge nearest the eye and β_F in minutes of arc of the edge farthest from the eye when the rectangle was at the most extreme slant used by that *E*. Columns 12 and 13 give the angular widths, δ , in minutes of arc of the nearest (subscript N) and farthest (subscript F) texture element from the eye. These angular widths (sizes) were measured at the meridian that bisects and is perpendicular to the axis of rotation when the surface was at the most extreme slant

used by that *E*. The regression coefficient is given as *b* (see Footnote 2 above).

Except for Freeman's and Epstein's experiments, which will be discussed later, $\beta_N - \beta_F$ for the cited experiments varied from 24 to 43 minutes of arc. Moreover, the angular sizes of the rectangles in these experiments varied between 6.3° and 9.8°. Consider the comparable data for the surface textures that were tested. For the cited experiments $\delta_N - \delta_F$ varied between 16.8 *seconds of arc* and 3.8 minutes of arc. Moreover, the angular sizes of the dot elements, δ , were never greater than 16.7 and were as small as 1.6 minutes of arc.

For these experiments, can we be sure that the angular changes in the texture elements over a slanted surface as well as the texture elements themselves were supraliminal for *Ss*? After all, angular changes of .28 minute-3.8 minutes of arc and even angular white and black dots of 1.6-16.7 minutes of arc would be expected to be subliminal under a variety of viewing conditions for even a normal monocular eye. Despite that, not a single *E* cited in Table 1 gave a solitary clue about the visual acuity of his *Ss*. Except for Smith (1964), not a single *E* who tested textures reported any luminance measures for dots and backgrounds. Smith (1964) reported the luminance of the white background but not of the black dots he used. It is not clear, therefore, that there was a discriminable change in the perspective of surface texture in any of these experiments.

That Freeman could have overlooked this crucial question is surprising. For one thing, not only did I discuss the question of surface texture and visual acuity, but I also predicted what Table 1 tends to reveal: that as a term like $\delta_N - \delta_F$ approaches liminal values, surface texture will be increasingly ineffectual in eliciting good slant judgments (Flock, 1964a, p. 386). Freeman's lapse here is surprising for a second reason. In his own work he attributed the failure of his first experiment to the use of overly small rectangles, even though for those rectangles $\beta_N - \beta_F$ was as large as 7.7 minutes of arc. He complained that "whatever cues were present in a ch

stimuli must have been very weak."³ According to Freeman there was no *discriminable* information in a perspective change of 7.7 minutes of arc when the stimulus object was a rectangle. Why, then, would he expect that there would be discriminable information when the stimulus object was a surface texture and the perspective change was as small as .28 minute–3.8 minutes of arc?

It is true, then, that in these studies surface texture has not elicited very accurate judgments of slant. But each instance of failure occurred under presumably near-liminal, liminal, or subliminal conditions, involving angularly small dot elements, angularly small changes in dot size, angularly small separations, low dot density, or short exposure and extreme variability of the elements of texture as in Flock's (1964b) Experiment IX. Experimental studies were available in which conditions were more favorable for surface texture, but Freeman seems to have overlooked them. For example, in the paper that he criticized (Flock, 1964a) there were at least three such references (Flock, 1962; Flock & Moscatelli, 1964; Gibson, 1950).

How effective is the linear perspective of a slanted, textureless, rectangular surface as a stimulus for slant? Is it as effective as Freeman implies? Except for Epstein's (1962), Epstein and Mountford's (1963), and Freeman's (in press) experiments, the untextured rectangular surfaces have been between 6.3° and 9.8° in height. For these rectangles all but one of the regression coefficients were between .30 and .44, the exception being Smith's (1964) coefficient of .67. Curiously, that result of .67 was greater than anything that Smith (1956, 1959) or his colleagues had found earlier, even though the experimental conditions were similar in most relevant respects. It would seem, therefore, that supraliminal 6°–10° rectangles at slants up to 40° and 50° will reliably yield a coefficient between .30 and perhaps .45. In contrast, *threshold* conditions for surface texture seem to yield coefficients between .22 and .30. Thus, under relatively optimal

conditions textureless rectangular surfaces perform only slightly better than do surface textures under threshold conditions.

FREEMAN'S AND EPSTEIN'S EXPERIMENTS

In Table 1 compare the results of Freeman and Epstein with all other *Es* who have tested the slant-inducing effects of textureless rectangular surfaces. Epstein and Freeman, using smaller rectangles than the other *Es* (see α in Table 1), get markedly better slant judgments. It is as if the smaller the rectangle, the better slant judgments became.⁴ There are some grounds, however, for disregarding the findings of both of these experimenters.

Freeman's (in press) data are, I believe, artifacts of his experimental method. Freeman predicted that if a standard (ST) slanted rectangle was smaller than a comparison (CO) rectangle, then when *S* adjusted the CO so that both would have the same perceived slant, *S* would underestimate the slant of the ST. Moreover, he predicted that the greater the discrepancy in size, the greater the underestimation. Conversely, if $ST > CO$, then *S* would overestimate the slant of the ST. Freeman's data from his second experiment purport to confirm this hypothesis. There are problems, however, in accepting this interpretation. Freeman presented ST and CO monocularly in two chambers of a tachistoscope, at time intervals of $ST = 1$ second, $rest = .7$ second, $CO = 1$ second. Both ST and CO appeared in the center of *S*'s field of view. *E* changed the slant of the CO in 2° steps between trials until, in effect, *S* reported they were at the same slant. It is not clear, however, that any *Ss* had to see rectangles at a slant in order to confirm Freeman's hypothesis. If CO was adjusted so that its projected trapezoidal image tended to be similar to that of the ST and so that the two were more or less centered in the field of view, one could predict all of Freeman's curves. In other words, *Ss*'

⁴ There is a nice irony here. Freeman's (in press) experimental thesis sought to prove that slant judgments became greater as rectangles were made *larger*.

³ The quoted explanation was made in an earlier edition (Freeman, 1964, p. 17).

tasks as they accepted them may have been to match trapezoids for similarity, rather than slanted rectangles for slant equality.

That Freeman's (in press) data are artifactual is indicated by his disclosure that some of his Ss in the first experiment reported that "the stimuli seldom appeared slanted to them." Nevertheless, 25 of the 27 regression coefficients for Ss in the first experiment lay between .83 and 1.17, the other two being .61 and .67. Since Ss said they could not see the slant of the rectangles, they must have produced these remarkably good results in some artifactual way as I suggested above. Moreover, if Ss could artifactually produce such remarkably good results in the first experiment, could they not artifactually produce the same good results in the second experiment? The second experiment differed from the first in that larger rectangles were used. The results of the second experiment were very similar to the first, 17 of 18 coefficients lying between .81 and 1.14, the 18th being .76.

There are other curiosities in Freeman's (in press) article. When the 1-centimeter rectangle was at 60° slant, he says that the angle subtended by the nearest edge minus the angle subtended by the farthest edge was 2.5 seconds of arc. Freeman says that this was "well below the limits of visual acuity"; and presumably, therefore, one would expect that Ss' judgments of slant should be 0°. In fact the *mean* judged slants for the 1-centimeter rectangle at 60° slant were 32°, 41°, 63°, and 58°, of which three are near approximations of the correct slant. Should not Freeman have explained how his Ss managed this remarkable feat? (When the rectangles were very small, Ss were told the direction of their slant and were both told and shown their true shape and size. Ss could then use the ratio of the angular projections of the slanted height to the unslanted width as a clue to the rectangle's slant.)

Epstein's (1962) and Epstein and Mountford's (1963) data should also not be taken too seriously. Epstein photographed a textureless rectangle, size unspecified, from which he made an Ekta-

chrome transparency, image size unspecified. The calculations in Table 1 are based on the assumption that, transparencies being about 1 inch, the image height on the transparency was at least 1 inch (although it might have been less). Since the rectangle was the size of a playing card, one might assume that the image width was .8 inch. On the basis of these assumptions $\beta_N - \beta_F$ for slants of 15°, 30°, 45°, and 60° was .20, .35, .55, and .65 minutes of arc, respectively. Despite these small perspective changes mean judged slants for the four slant levels were 15°, 30°, 33°, and 44° (Epstein, 1962) and 20°, 29°, 31°, and 54° (Epstein & Mountford, 1963). These performances were so remarkably good, one wonders whether artifacts were not present in these experiments also.

On grounds other than these, however, it may be desirable not to take Epstein's data seriously. For part of his 1963 study with Mountford he used trapezoids in a frontal plane as STs, comparing the slant judgments they induced with those induced by physically slanting the transparencies. His trapezoids were wholly incommensurate with the slanted rectangles, however. The trapezoids were projections of much more extremely slanted rectangles than of the slanted rectangles with which they were compared. Although he did refer to these as "slight" discrepancies, that fact did not dissuade him from concluding that frontal trapezoids induced better slant judgments than the slanted rectangles of which they were purported to be the projection. Nor did it persuade him to discard the data and start over again. (Smith, 1964, for example, attempted and failed to replicate his findings.)

GRUBER AND CLARK'S EXPERIMENT

Gruber and Clark's (1956) experiment (see above and Table 1 for details) purports to be far more critical of my theoretical position than for the reason given by Freeman (1965, p. 503). Gruber and Clark report only part of their data, in the form of four curves (their Figure 2). The relative position of these curves acquired great significance for Gruber

and Clark and later for Eriksson (1964).⁵ Gruber and Clark theorized that judged slant depended on a cortical or retinal interaction unrelated to measures of visual acuity. They predicted that dots which are close together or far apart will yield slant judgments that are worse than for dots which are at interim separations. Moreover, they predicted that there would be increasingly accurate slant judgments as one went from extremely small or extremely large to moderate separations. In their second experiment they tested and confirmed these predictions.

Consider the facts, however. Except for their viewing distance of 1.5 meters the total fluctuation over all their mean judged slants was apparently less than 6.5° (extrapolating from their Figure 2). This range of 6.5° determined the relative positions of their four curves. Their display of 3-millimeter dots with 8-millimeter separations at 1.5 meters yielded an apparent mean judged slant of 10.1° , whereas their display of 6-millimeter dots with 16-millimeter separations at 3.0 meters yielded an apparent mean judged slant of 14.5° , a difference of 4.4° . These two displays were *optically identical*. If under the same conditions the same Ss vary 4.4° , can one seriously entertain interpretations of four curves based on a total fluctuation among them of 6.5° ?

In their second experiment they constructed three surfaces each textured with round, white 6-millimeter dots. The separation of the dots measured from center to center for the three surfaces was 6, 30, and 90 millimeters. A 6-millimeter dot separated by 6-millimeters means that the dots had to be adjacent to each other, leaving an interspace where the dots met of approximately 1.5 millimeters in height and shaped like an equilateral triangle. At the viewing distance of 1,380 millimeters with a field of view of 15.9° and a single surface slant of 43° , the angular length of that interspace at places farthest from the eye was about 2 minutes of arc. Mean judged slant was 10.4° . Is it not possible that, for some Ss over presumably the 48 trials of the experiment, visual acuity for so small an interspace

could have sometimes been subliminal? For some Ss on some trials perhaps the farthest part of the surface, perhaps all of the surface, appeared homogeneous and untextured. Gruber and Clark do not report any data at all about the visual acuity of their Ss. They do not report illuminance or dot-background contrast. I would predict that judgments of slant ought to be poor under their conditions. At the other extreme, when dot separation was 90 millimeters, there were 20 dots over the entire display, and mean judged slant was 6.7° . For the interim condition when dot separation was 30 millimeters there were 180 elements over the display, and judged slant was 13.5° . In other words, with an increase in density from 20 to 180, judged slant improved, as my Equation 4 predicts (Flock, 1964a, p. 383). When dot elements were made so dense that for some Ss parts of the surface might have appeared untextured, perceived slant decreased. One hardly need postulate a theory of cortical or retinal interaction to explain that result. Moreover, if their interaction hypothesis depends on the extreme conditions of their second experiment, its relevance to optical variables of texture is debatable.

LIMITATIONS OF CONTOUR CONVERGENCE

Freeman's reluctance to accept any other variable than the contour convergence of slanted rectangles as a stimulus for perceived slant is hard to reconcile with the facts of man's natural environment. In nature, apart from civilized structures, there seem to be few rectangles and rectangular-like elements. Nevertheless, animals and men (in the jungle, for example) make responses that are neatly attuned to changes in the slants of things, just as if they were correctly perceiving slants. Correspondingly, how does one use linear perspective to explain the following unpublished experiment by Flock and Graves?

An extended surface was constructed with 518 different-sized, different-shaped triangles distributed randomly over it. All surface parameters were normally distributed, and the angular mean height and σ of the distribution of shapes at the

⁵ Most of the criticisms leveled against Gruber and Clark apply equally to Eriksson.

center of the display when the surface was at 0° slant were 4.5° and 2.33° , respectively. The surface was placed at six different slants and was seen through apertures of 20° , 40° , and 80° . Regression coefficients for the three fields of view with 4.3, 19, and 126 shapes visible at 30° slant were .10, .30, and .69, respectively. A second group of Ss judged the slant of a small (14.03°), medium (32.25°), and large (66.25°) triangle under the same conditions and produced coefficients of .09, .03, and $-.04$, respectively. Thus, contours of triangles induced no slant regardless of their size, whereas textures composed of small triangles induced slant judgments that became progressively better as field of view was increased. Can Freeman explain these results in terms of contour convergence? It is doubtful that he can explain any slant data in which a rectangle is not presented. His unqualified claim that contour convergence is *the* stimulus for perceived slant is both oversimplified, because it does not handle all of the available data, and premature, because it is even less capable of coping with the kind of data just presented.

THE SPECIAL CASE OF THE RECTANGLE

Nowhere have I asserted that the optics of surface texture constitute *the* stimuli for slant or are the sole determiners of its perception. I have never made any claim for variables of optical texture other than to show that they *could* explain some old problems in perception. For example, given variables of optical texture, it is not necessary to introduce into space perception many of the subjectivist assumptions with which we have been burdened (see also Flock, 1964c). The theoretical model I have proposed makes it possible to give an account of how perspective cues like linear perspective, size perspective, and motion parallax are related. Freeman's (1965) paper does raise the question, however, of how and whether the isolated slanted rectangle is related to my theoretical formulation. The answer is that the rectangle can be considered an extreme and special case of the formula-

tion.⁶ There are many aspects of perspective, however, that the theoretical model will not handle. That fact does not mean that either the theoretical model or some event it will not handle is *bad*, as is implied by Clark et al. (1956b) and Smith (1956) whose taxonomy refers to the elliptical projection of a slanted circle as a "distortion." What it does mean is that the theoretical model has limitations.

In the theoretical model the idea of redundancy of optical information was developed. It was shown that there would be redundant classes of congruent luminous patterns, any one of which would theoretically specify a distal surface slant. These redundant sources of information might refer to a distal rectangle, to the surface texture of that rectangle, or to a specific structural aspect of that surface texture. What makes a particular source of information an effective rather than a potential stimulus for a visual system could depend on a variety of factors, for example, viewing conditions, visual resolving powers, a variety of physiological conditions, instructions, attentional factors, an endless variety of experiential factors, etc. (see Flock, 1963, 1964a; Flock & Moscatelli, 1964; and for the distinction between potential and effective stimuli, see Gibson, 1960). For these reasons, pitting the force of one class of perspective information against the force of a second, as is favored by many *Es* cited above, might

⁶ With the optic fixation point at the center of the projected trapezoid, with the normal meridian angularly equidistant between the parallel trapezoidal bases and approximately projecting the true axis of rotation, with the great circle arc through the optic fixation point and perpendicular to the normal meridian, it will then be true along the normal meridian and all meridians parallel to it that (a) angular extents of the projected trapezoid will be bilaterally symmetrical on either side of the great circle arc, (b) there will be a gradient of those angular extents along the great circle arc, and (c) therefore Criteria I-IV will be met, although the conditions are special. The slant of the rectangle is then given as specified by the derivation of optical *theta* (see Flock, 1964a, pp. 384-388).

yield surprisingly little useful information about the *general* effectiveness of a particular stimulus for space perceptions.

PHYSIOLOGICAL EVIDENCE

Freeman finds in several physiological studies support for contour convergence as the "principal stimulus" for slant. The reasoning (1965, p. 503) is that the "visual system is 'tuned' to register contours (abrupt brightness gradients)." Freeman seems to equate linear perspective with all contours and with all brightness gradients. In fact, my lengthy discussion of congruent luminous patterns *depends* on the assumption that visual systems *are* tuned to register contours and abrupt brightness gradients (Flock, 1964a, pp. 380-382). How else could the optical texture be registered by the organism?

FLOCK'S EQUATION 4

Freeman (1965, Footnote 2) attributes to this equation peculiar properties, of which negative values is one. But the equation cannot take negative values. The equation depends on the assumption that the parameters of a texture are, if not normally, at least rectangularly distributed (Flock, 1964a, p. 382). Under those conditions it cannot have negative values. Even without such restriction, the conditions under which it could take negative values are very extreme.

Equation 4 should not be considered more than a first statement of the relation between size of visual field, size of element, and variability of elements, on the one hand, and stochastic regularity from the point of view of optics, on the other. Its function at present is a strategic one. Perhaps for the first time three extremely important optical variables have been brought together into a rational relationship. Until the effects of these variables have been empirically tested, their precise relationship must remain unknown.

FLOCK'S EQUATION 5

According to Freeman (1965, p. 502) "Flock's optical theory of slant encounters its greatest trouble" when it relates the optical slant of the surface to ratios of angles, rather than to the surface's "size, distance, and other variables, as well as

slant." He then cites some commonplaces about perspective and implies that they make invalid my Equation 5. He is incorrect. Equation 5 is invariant over the changes in conditions which he describes. Freeman is simply wrong when he attempts to discuss the mathematics of perspective.⁷ Finally, his complaint that the derivation of Equation 5 was not available is unwarranted (Flock, 1964a, Footnote 6).

SENSORY ABILITIES

Freeman raises questions about the sensory abilities presumed to be necessary for the registration of the optical variables that specify the slant of a surface. Most of the presumed abilities are not very extraordinary. I do want to be explicit about one visual task, however, that is crucial to the assumption that these potential optical variables might be perceptually effective. In the section on "Magnification" I discussed the luminous structure of the optical texture. I suggested that an optical instrument like the eye could scan a small region of the optical texture in some angularly uniform step, register changes in absolute (but averaged) luminances for successive steps, adjust the angular magnitude of the step in scanning a different region of the optical texture, and thereby identify congruent luminous patterns over the optical texture. Whether an eye could make this kind of optical analysis is, of course, an empirical question. Pursuing that empirical question is only one of the

⁷ Consider Freeman's (in press) use of mathematics in his own work. That his Equation 2 is false should be self-evident. For example, letting a and d have reasonable values and letting $c \rightarrow 0$, $\tan \delta$ should approach 0. His $\tan \delta$ does not have this property. His statement that π is unchanged by the height of the rectangle is self-evident but has absolutely nothing to do with the stimulus at the eye or with his argument. The angular size of π at the eye very definitely changes with the height of the rectangle. He then compares π with δ , the former being a base angle on a projection plane, the latter being the angular difference of a different event at the eye. For a first-rate analysis of the complexities of perspective, see Gibson (1957).

possible strategies, however. I have recently been fascinated by the many-sided hypotheses that can be generated from the theoretical model and by the possibility of testing such hypotheses. For example, applying Jameson and Hurvich's (1964) achromatic color theory to the visual problem of identifying congruent luminous patterns in the optical texture, it seems to me that it might be possible to predict systematic errors in slant and distance perception of extended surfaces as a function of changing illuminance and luminance levels. Hypotheses that have been tested and are in the process of being tested have been discussed elsewhere (Flock, 1964a). The data from these tests have not been published as yet, however.

CONCLUSION

Freeman calls this theoretical model a theory of the stimulus and calls his own a theory of perception. I do not know what is gained by that distinction. Presumably, theories that account for and predict empirical results will flourish, the others will not.

REFERENCES

- CLARK, W. C., SMITH, A. H., & RABE, A. Retinal gradients of outline as a stimulus for slant. *Canadian Journal of Psychology*, 1955, 9, 247-253.
- CLARK, W. C., SMITH, A. H., & RABE, A. The interaction of surface texture, outline gradient, and ground in the perception of slant. *Canadian Journal of Psychology*, 1956, 10, 1-8. (a)
- CLARK, W. C., SMITH, A. H., & RABE, A. Retinal gradients of outline distortion and binocular disparity as stimuli for slant. *Canadian Journal of Psychology*, 1956, 10, 77-81. (b)
- DEMBER, W. N. *The psychology of perception*. New York: Holt, Rinehart, & Winston, 1961.
- EPSTEIN, W. Apparent shape of a meaningful representational form. *Perceptual and Motor Skills*, 1962, 15, 239-246.
- EPSTEIN, W., & MOUNTFORD, D. Judgment of slant in response to an isolated gradient of stimulation. *Perceptual and Motor Skills*, 1963, 16, 733-737.
- ERIKSSON, E. S. Monocular slant perception and the texture gradient concept. *Scandinavian Journal of Psychology*, 1964, 5, 123-128.
- FLOCK, H. R. *The monocular perception of surface slant*. (Doctoral dissertation, Cornell University) Ann Arbor, Mich.: University Microfilms, 1962, No. 62-2514.
- FLOCK, H. R. Selective registration of information in visual judgments of surface slant. *Perceptual and Motor Skills*, 1963, 17, 537-538.
- FLOCK, H. R. A possible optical basis for monocular slant perception. *Psychological Review*, 1964, 71, 380-391. (a)
- FLOCK, H. R. Some conditions sufficient for accurate monocular perceptions of moving surface slants. *Journal of Experimental Psychology*, 1964, 67, 560-572. (b)
- FLOCK, H. R. Three theoretical views of slant perception. *Psychological Bulletin*, 1964, 62, 110-121. (c)
- FLOCK, H. R., & MOSCATELLI, A. Variables of surface texture and accuracy of space perceptions. *Perceptual and Motor Skills*, 1964, 19, 327-334.
- FREEMAN, R. B., JR. The effect of size on judgments of slant. Research Bulletin No. 45, 1964, Department of Psychology, Pennsylvania State University.
- FREEMAN, R. B., JR. Ecological optics and visual slant. *Psychological Review*, 1965, 72, 501-504.
- FREEMAN, R. B., JR. Effect of size on visual slant. *Journal of Experimental Psychology*, 1966, 71, in press.
- GIBSON, J. J. The perception of visual surfaces. *American Journal of Psychology*, 1950, 63, 367-384.
- GIBSON, J. J. Optical motions and transformations as stimuli for visual perception. *Psychological Review*, 1957, 64, 288-295.
- GIBSON, J. J. The concept of the stimulus in psychology. *American Psychologist*, 1960, 15, 694-703.
- GRUBER, H. E., & CLARK, W. C. Perception of slanted surfaces. *Perceptual and Motor Skills*, 1956, 6, 97-106.
- JAMESON, D., & HURVICH, L. Theory of brightness and color contrast in human vision. *Vision Research*, 1964, 4, 133-154.
- SMITH, A. H. Gradients of outline convergence and distortion as stimuli for slant. *Canadian Journal of Psychology*, 1956, 10, 211-218.
- SMITH, A. H. Outline convergence versus closure in the perception of slant. *Perceptual and Motor Skills*, 1959, 9, 259-266.
- SMITH, A. H. Judgment of slant with constant outline convergence and variable surface texture gradient. *Perceptual and Motor Skills*, 1964, 18, 869-875.

(Received April 15, 1965)

